

Dimensionen politischer Konflikte: die Analyse von namentlichen Abstimmungen in Parlamenten mit dem Verfahren der Mokken-Skalierung

Best, Heinrich; Zimmermann, Ernst Joseph

Veröffentlichungsversion / Published Version

Sammelwerksbeitrag / collection article

Empfohlene Zitierung / Suggested Citation:

Best, H., & Zimmermann, E. J. (1991). Dimensionen politischer Konflikte: die Analyse von namentlichen Abstimmungen in Parlamenten mit dem Verfahren der Mokken-Skalierung. In H. Best, & H. Thome (Hrsg.), *Neue Methoden der Analyse historischer Daten* (S. 41-79). Sankt Katharinen: Scripta Mercaturae Verl. <https://nbn-resolving.org/urn:nbn:de:0168-ssoar-338081>

Nutzungsbedingungen:

Dieser Text wird unter einer Deposit-Lizenz (Keine Weiterverbreitung - keine Bearbeitung) zur Verfügung gestellt. Gewährt wird ein nicht exklusives, nicht übertragbares, persönliches und beschränktes Recht auf Nutzung dieses Dokuments. Dieses Dokument ist ausschließlich für den persönlichen, nicht-kommerziellen Gebrauch bestimmt. Auf sämtlichen Kopien dieses Dokuments müssen alle Urheberrechtshinweise und sonstigen Hinweise auf gesetzlichen Schutz beibehalten werden. Sie dürfen dieses Dokument nicht in irgendeiner Weise abändern, noch dürfen Sie dieses Dokument für öffentliche oder kommerzielle Zwecke vervielfältigen, öffentlich ausstellen, aufführen, vertreiben oder anderweitig nutzen.

Mit der Verwendung dieses Dokuments erkennen Sie die Nutzungsbedingungen an.

Terms of use:

This document is made available under Deposit Licence (No Redistribution - no modifications). We grant a non-exclusive, non-transferable, individual and limited right to using this document. This document is solely intended for your personal, non-commercial use. All of the copies of this documents must retain all copyright information and other information regarding legal protection. You are not allowed to alter this document in any way, to copy it for public or commercial purposes, to exhibit the document in public, to perform, distribute or otherwise use the document in public.

By using this particular document, you accept the above-stated conditions of use.

Dimensionen politischer Konflikte

Die Analyse von namentlichen Abstimmungen in Parlamenten mit dem Verfahren der Mokken-Skalierung

Heinrich Best u. Ernst J. Zimmermann

1. Methoden und Modelle der Analyse namentlicher Abstimmungen

Es ist ein häufig beschworenes Dilemma der quantifizierenden historischen Sozialforschung, daß sie auf "vorgefundene" Daten angewiesen ist und ihre Indikatoren typisch nicht aus direkten, durch den Forscher selbst kontrollierten Meßoperationen gewinnen kann. Diese Beschränkung hat weitreichende Konsequenzen. So werden Individualdaten in der historischen Sozialforschung häufig auf niedrigem Meßniveau erhoben, die von ihr ausgewerteten Massenquellen sind weit überwiegend Erzeugnisse "öffentlicher Buchführung", die eher institutionell definierte Rollen und formale Strukturen als die intentionalen und motivationalen Aspekte menschlichen Handelns abzubilden vermag - Ausnahmen bestätigen auch hier die Regel.¹

Das ist eine fatale Diagnose, und es ist verständlich, warum historische Sozialforscher mit erheblichem Einsatz bemüht sind, das hermetische Gehäuse formaler Strukturanalysen zu sprengen. Die Auswertung der mentalitätsgeschichtlich ertragreichen Massenquellen oder die Anwendung von Verfahren computerunterstützter Inhaltsanalyse auf große Textmengen eröffnen Auswege. Seltener werden statistische Modelle angewendet, die hinter der Fassade manifester Beobachtungswerte verborgene "latente Strukturen" und "konzeptuelle Schemata" historischer Akteure aufzudecken vermögen. Zu diesem Verfahren zählen neben anderen die "intrinsischen" oder "indirekten" Skalierungsmethoden, die "numerische Strukturen, gemäß modellspezifischen Annahmen, in andere numerische Strukturen transformieren".² Sie sind insbesondere dann anwendbar, wenn mangels eines externen Kriteriums die Gültigkeit der Messung eines latenten Konstrukts nur aus den Wechselbeziehungen zwischen den Indikatoren selbst erschlossen werden kann.

Die Voraussetzung für den Einsatz indirekter Skalierungsverfahren ist die

¹ Vgl. hierzu u.a. die Zusammenfassung der "qualitativen Kritik" und der "quantitativen Selbstkritik" in Konrad H. Jarausch et al., Quantitative Methoden in der Geschichtswissenschaft (Darmstadt 1985), S. 195-203.

² Erwin K. Scheuch u. Helmut Zehnpfennig, "Skalierungsverfahren in der Sozialforschung", in: Handbuch der empirischen Sozialforschung, Bd. 3a (3. Aufl., Stuttgart 1974), S. 99.

Überlieferung möglichst umfangreicher Serien gleichartiger "Indikatoren", "Stimuli" oder "Items". Sie ist in historischen Untersuchungen häufiger gegeben als gemeinhin angenommen wird. Ein besonders wichtiges Anwendungsgebiet ist das formale Abstimmungsverhalten in Organisationen, vor allem in rechtsprechenden und legislativen Körperschaften.³ Wird das Entscheidungsverhalten der einzelnen Mitglieder dabei in allen, zumindest aber in einer größeren Anzahl von Abstimmungen protokolliert und werden diese Voten nicht vollständig durch feste Fraktions- und Koalitionsbindungen prädeteminiert, dann ist es sinnvoll, nach "latenten" kognitiven oder normativen Konzepten zu suchen, die das Abstimmungsverhalten strukturierten. Tatsächlich zeigt sich in einschlägigen Untersuchungen regelmäßig, daß sich die beteiligten Akteure in wiederkehrenden Abstimmungssituationen nicht jedesmal "ad hoc" entscheiden, sondern mehr oder weniger prägnanten Abstimmungsmustern folgen. Die Frage, worin diese Regelmäßigkeiten begründet sind, wird dann allerdings sehr unterschiedlich und zum Teil kontrovers beantwortet.

Vor allem in den USA, Großbritannien, den Niederlanden und den skandinavischen Ländern war die Analyse namentlicher Abstimmungen unter der Bezeichnung "roll-call analysis" eines der ersten und ist nach wie vor eines der wichtigsten Anwendungsgebiete quantitativer Methoden in der Geschichtswissenschaft und in der historischen Soziologie. Sie wird eingesetzt, um die grundlegenden politischen Konfliktlinien in Parlamenten aufzudecken, den Einfluß externer Akteure auf legislative Entscheidungen zu identifizieren, den "Lebenszyklus" ideologischer Konflikte zu verfolgen und Prozesse der Koalitions- und Fraktionsbildung zu rekonstruieren.⁴ Die Bedeutung serieller Abstimmungsanalysen für die quantifizierende Geschichtsforschung hat William O. Aydelotte mit geradezu emphatischen Worten hervorgehoben. Seine sehr optimistische Bewertung des empirischen Gehalts und der Qualität von Abstimmungsdaten ist aber auch ein Beleg für die gelegentliche Naivität, in der sie vor allem in der Frühphase nach ihrer "Entdeckung" als Massenquelle verwendet wurden:⁵ "Yet there exists a source, a body of materials, ... which is

³ Vgl. u.a. Melissa P. Collie, "Voting Behaviour in Legislatures", in: Gerhard Loewenberg et al. (Hrsg.), *Handbook of Legislative Research* (Cambridge/Mass. u. London 1985), S. 471-518. Darin auch weitere Hinweise auf die umfangreiche amerikanische Literatur zur legislative behavior-Forschung.

⁴ Aage R. Clausen, "Some Basic Approaches to the Measurement of Roll-Call Voting", in: Donald F. Freeman (Hrsg.), *Foundation of Political Science* (New York u. London 1977), S. 313.

⁵ "Voting Patterns in the British House of Commons in the 1840's", in: *Comparative Studies in Society and History*, 5 (1962-1963), S. 134f.

not only readily available but is also fuller and more reliable than most sources which historians have at their disposal. This is the division lists. ... The peculiar value of these lists is that they contain expressions of opinion upon important subjects from men who are no longer available for questioning. It does not strain terms too much to say that, imaginatively used, they can constitute a kind of questionnaire which one may submit to the dead on which they will give us their replies. Further, the information about the votes of members of parliament is perhaps more complete and more certain than anything else that is known about them. The social background, their economic interests and their relations with their constituents can be studied with profit, matters the information is and always will be fragmentary and there are nuances which it is difficult now to recapture. By contrast, the stands which those men publicly adopted on the major issues of the day are documented by the division lists with a wealth of detail, repeated corroboration and, reason to believe, a relatively high degree of accuracy".

Es ist bemerkenswert, daß Verfahren der seriellen Abstimmungsanalyse in der Bundesrepublik Deutschland bislang noch kaum eingesetzt wurden. Die Gründe dafür können hier nicht ausführlich erörtert werden. Sie sind jedenfalls nicht in der geringeren Dichte der Überlieferung oder der minderen Relevanz von namentlichen Abstimmungen in den historischen Parlamenten West- und Mitteleuropas zu suchen. Ausschlaggebend war wohl der Umstand, daß angesichts der vor allem in der heutigen Bundesrepublik Deutschland hochdisziplinierten Parlamente kein Ausstrahlungseffekt von der gegenwartsbezogenen Forschung ausgeht, der historische Untersuchungen mit Abstimmungsdaten hätte stimulieren können.⁶ In dieser Hinsicht unterscheidet sich die Forschungssituation in der Tat grundlegend von der in den USA oder anderer europäischer Länder, in denen seit langem in historischen und gegenwartsbezogenen Untersuchungen Abstimmungsanalysen erfolgreich angewendet werden. Die sozialwissenschaftliche Relevanz eines Untersuchungsgegenstandes bestimmt sich jedoch nicht an seinem Gegenwartsbezug, sondern an seinem Gehalt verallgemeinerungsfähiger Erkenntnisse - wobei für Historiker selbst ohne solche Generalisierungsabsichten das Abgelebte ein würdiger Gegenstand des Interesses ist.

Im folgenden soll nun eine Forschungssituation vorgestellt werden, in der sich mit Abstimmungsanalysen Prozesse der Strukturbildung in Parlamenten

⁶ Ausführlich zur Forschungssituation und zur Methodik der Abstimmungsanalyse jetzt Heinrich Best, *Die Männer von Bildung und Besitz. Struktur und Handeln parlamentarischer Führungsgruppen in Deutschland und Frankreich 1848/49* (Düsseldorf 1990), S.484-531.

untersuchen lassen: "die Stunde Null" nach der Ausdehnung der Partizipationsrechte infolge von Revolutionen. Solche Versammlungen werden als Interaktionssysteme betrachtet, die noch ganz am Anfang ihrer Systemgeschichte stehen, in der sich noch wenig an systemspezifischen Institutionalisierungen, Erfahrungen und Erinnerungen, oder, um es mit Luhmann zu sagen: an "aufbewahrter Selektivität", abgelagert hat.⁷ Das Jahr 1848 war eine solche "Stunde Null", in der sich zum ersten Mal in Mittel- und Westeuropa das allgemeine Männerwahlrecht mit parlamentarischer Regierungsweise, einer Formierung von Parlamentsfraktionen, einer breiten politischen Mobilisierung der Bevölkerung und Anfängen organisierter Massenparteien verband. Auch die Abgeordneten der beiden konstituierenden Versammlungen, die im Mai 1848 in Paris und Frankfurt zusammentraten, waren, was die Parlamentserfahrung angeht, weit überwiegend homines novi: Nur jeweils etwa 20% der Mitglieder beider Konstituanten hatten schon vor dem Mai 1848 in einer Kammer oder einem Landtag gesessen.⁸

Hier interessiert nun vor allem die soziale Morphologie der Konfliktstrukturen, die sich in der Situation relativer Freiheit eines institutionellen Neubeginns ausdifferenzierten. Dazu wählten wir das Verfahren der seriellen Analyse namentlicher Abstimmungen, für das in beiden Untersuchungsgruppen die Anwendungsbedingungen günstig sind. So wurde in der Pariser Konstituante insgesamt 259 mal abgestimmt. Die Geschäftsordnung sah dies vor, wenn mindestens zwanzig Abgeordnete einen entsprechenden Antrag unterstützten. Obligatorisch waren namentliche Abstimmungen vor allem bei unmittelbar ausgabenwirksamen Beschlüssen. In der Frankfurter Nationalversammlung erfolgte eine offene Abstimmung durch Namensaufruf, wenn dies am Schluß der Debatte von mindestens fünfzig Abgeordneten verlangt wurde. Zwar war das Quorum hier strenger als in der Pariser Konstituante, doch wurde in der Paulskirche bei einer kaum längeren Sitzungsperiode noch häufiger - über insgesamt 298 Vorlagen - namentlich abgestimmt. Wie im Fall der französischen Nationalversammlung wurden die Voten der Abgeordneten protokolliert und anschließend in den offiziellen Sitzungsprotokollen publiziert. Insgesamt liefert das namentliche Abstimmungsverhalten für beide Untersuchungsgruppen mehr als 470000 Datenpunkte zum parlamentarischen Entscheidungsverfahren der Abgeordneten.

⁷ "Weltzeit und Systemgeschichte. Über Beziehungen zwischen Zeithorizonten und sozialen Strukturen gesellschaftlicher Systeme", in: Peter Christian Ludz (Hrsg.), Soziologie und Sozialgeschichte (Opladen 1973), S. 85.

⁸ Best, Bildung und Besitz, S. 170.

Diese Fülle ist zwar eine besonders beeindruckende Widerlegung des Vorurteils von der relativen Datenarmut der historischen Sozialforschung, doch wäre es vorschnell zu glauben, daß man hier über einen unproblematischen Indikator verfügt, der einen ungebrochenen empirischen Zugang zu einer Analyse der politischen Präferenzen der Abgeordneten eröffnet. Tatsächlich sind namentliche Abstimmungen der Ausdruck eines hochkontingenten sozialen Handelns, in dem die ideologischen Präferenzen der Abgeordneten nur ein Bestimmungsmoment unter anderen sind. Abgeordnete sind nutzenmaximierende Akteure, die unter verfügbaren strategischen Alternativen wählen, um ihre Ziele zu erreichen. Junktimpositionen und Obstruktion sind Beispiele für taktisches Verhalten, die nahelegen, daß Abweichungen von der persönlichen Präferenzskala häufige Vorkommnisse in parlamentarischen Abstimmungen sind.⁹ Solchen Situationen ist gemeinsam, daß Abgeordnete einem Entscheidungskalkül folgen, das andere Akteure in ihrem Umfeld einbezieht, die ebenfalls unmittelbaren Einfluß auf das Ergebnis der Abstimmung haben. Deshalb gilt: während die Folgenlosigkeit der Antworten für die Befragten die Voraussetzung für die Gültigkeit von Ergebnissen in der Standard-Interviewsituation ist, sind die politische Wirkung und die absichtsvolle Sichtbarkeit des Abstimmungsverhaltens Kennzeichen von Entscheidungssituationen in parlamentarischen Versammlungen.

Für die Analyse des namentlichen Abstimmungsverhaltens resultiert hieraus das schwer lösbare Problem, die "strategische Komponente" in den Voten ermitteln zu müssen, wenn man aus ihnen Indikatoren für die politische Orientierung der Abgeordneten gewinnen will. Ein solches Vorgehen setzt die Konstruktion eines allgemeinen entscheidungstheoretischen Modells voraus, auf dessen Grundlage das Stimmverhalten von Abgeordneten unter der Voraussetzung vorhergesagt werden kann, daß sie entsprechend ihren persönlichen Präferenzskalen votieren. Abweichungen zwischen dem danach erwarteten Abstimmungsmuster und dem tatsächlich beobachteten Abstimmungsverhalten lassen dann auf eine strategische Komponente schließen. Erst in den vergangenen Jahren wurden Modelle parlamentarischen Abstimmungsverhaltens entwickelt, die die Elemente von Entscheidungssituationen und deren konditionelle Verknüpfung soweit spezifizieren, daß ein solches Vorgehen aussichtsreich wird.¹⁰ Zuvor waren Abstimmungsanalysen zumeist kaum mehr als beliebige Kombinationen eines Datentyps mit einem Methodenre-

⁹ J.A. Frerejohn u. M.P. Fiorina, "Purpose Models of Legislative Voting", in: *American Economic Review*, 65 (1975), S. 407.

¹⁰ Vgl. insbes. Menno Wolters, *Interspace Politics* (Leiden 1984).

pertoire, das nach Vorkenntnis und Vorliebe des einzelnen Forschers eingesetzt wurde.

Ausgehend von dem entscheidungstheoretischen Axiom, daß das zentrale Element einer Entscheidung die Bewertung eines Items nach einem Kriterium ist, wird Abstimmungsverhalten in Analogie zu Ansätzen moderner Theorien des Wählens als ein mehrdimensionales räumliches Modell konzipiert. Die Grundform des Kalküls bei einer Abstimmung ist danach die Bewertung der Distanz zwischen dem "idealen Punkt" des Abgeordneten auf seiner Präferenzskala und dem Punkt, den er einer Abstimmungsvorlage zuordnet. Die gerade noch tolerierte Abweichung vom Idealpunkt ist die kritische Distanz.

Während bei dem so modellierten Bewertungsprozeß die Entscheidung für oder gegen eine Vorlage ein interner Vorgang ist, wird bei der Berücksichtigung einer expliziten Alternative ein externes (kognitives) Kriterium eingeführt. Herbert Weisberg, auf den die hier skizzierte Unterscheidung zurückgeht, spricht von einem "cumulative scale model", wenn eine Alternative bei der Entscheidung berücksichtigt wird, von einem "proximity scale model", wenn dies nicht der Fall ist.¹¹ Der Unterschied zwischen beiden Entscheidungsmodellen ist insofern empirisch bedeutsam, als "proximity voting" die Möglichkeit einschließt, daß die Abgeordneten auf beiden Flügelpositionen übereinstimmend alle Abstimmungsvorlagen ablehnen, weil sie außerhalb des Bereichs ihrer kritischen Distanzen liegen. Dies ist dann der Fall, wenn die Abstimmungsgegenstände im mittleren Bereich des politischen Kontinuums lokalisiert sind, die Abgeordneten aber über die gesamte Spannweite des Konfliktraumes streuen. Das hat mißliche Folgen, wenn die politischen Positionen von Abgeordneten durch ihr Abstimmungsverhalten bestimmt werden sollen: "Les extrêmes se touchent", die Flügelpositionen werden sich dann nicht unterscheiden lassen, selbst dann, wenn alle Abgeordneten bei jeder Abstimmung ihren persönlichen Präferenzen entsprechend votieren und strategisches Abstimmungsverhalten nicht ins Spiel kommt. Wenn sich Abgeordnete zwischen expliziten oder impliziten Alternativen entscheiden, ist eine solche Konstellation hingegen ausgeschlossen.

Der Status quo ist die in der parlamentarischen Praxis häufigste, wenn auch fast immer implizite Alternative bei Abstimmungen. Bleibt der Status quo über Serien von Voten die relevante Entscheidungsalternative und markiert er zugleich einen der beiden Pole der Präferenzskalen aller Abgeord-

¹¹ "Scaling Models for Legislative Voting", in: American Political Science Review, 66 (1972), S. 1306-1315.

neten, ist der Sonderfall der "Unipolarität" gegeben. Das ist ein Umstand, der das Problem der Ermittlung von politischen Orientierungen durch Abstimmungsanalysen in entscheidender Weise vereinfacht. Denn unter der weiteren Voraussetzung, daß Abgeordnete und Abstimmungsgegenstände über den gleichen Bereich des politischen Kontinuums streuen - wofür in unserer Untersuchungssituation die Beobachtung spricht, daß die Vorlagen von den Abgeordneten selbst formuliert und eingebracht wurden - lassen sich dann Abstimmungen und Abgeordnete in eine korrespondierende Rangordnung bringen. Die resultierende Skala hätte idealiter die Eigenschaften einer Guttman-Skala.¹² Deren Meßmodell wurde von Stouffer prägnant beschrieben: "Es muß möglich sein, die einzelnen Indikatoren so zu ordnen, daß im Idealfall alle Personen, die eine gegebene Frage (hier: eine gegebene Abstimmungsvorlage) zustimmend beantworten, höhere Ränge einnehmen als Personen, die die gleiche Frage ablehnend beantworten. Aus dem Rang oder dem Skalenprodukt einer Person können wir exakt schließen, welche Indikatoren sie bevorzugten".¹³

In unserer Untersuchungssituation stellt sich nun die Frage, ob sich die namentlichen Abstimmungen einer oder mehreren kumulativen Skalen zuordnen lassen und wie sich die Abgeordneten auf die Skalenpositionen verteilen. Als Verfahren zur Rekonstruktion der parlamentarischen Konfliktstrukturen haben wir die Mokken-Skalierung eingesetzt. Es ist dies eine stochastische Version des Guttman-Modells, das ja von der rigorosen Annahme einer deterministischen Relation zwischen der empirisch zugänglichen Beobachtungsebene und der Konstruktebene ausgeht. Das ist eine wenig realistische Annahme, und Guttman-Skalen, die dem Kriterium folgen, daß einem Beobachtungspunkt ein und nur ein Punkt auf einem latenten Kontinuum entspricht, sind in der empirischen Sozialforschung dann auch extrem selten. "Latent-structure" Modelle wie die Mokken-Skalierung gehen demgegenüber davon aus, daß für die Abweichungen zwischen dem Guttman-Modell und der beobachteten Realität außer Mehrdimensionalität auch unscharfe Valenzen der Items verantwortlich sein können.¹⁴ Danach besteht zwischen einem Indikator und einer latenten Einstellungsdimension eine probabilistische Beziehung: Der Indikator wird nicht immer und unter allen Bedingungen die gleiche Information über das Konzept liefern. Die Mokken-

¹² In M. Wolters verzweigter Systematik von Abstimmungsmodellen entspricht die Guttman-Skala dem "unipolar unidimensional individualized fixed alternative model". Interspace Politics, S. 50.

¹³ Samuel A. Stouffer, *Measurement and Prediction*, Princeton 1950, S. 5,

¹⁴ C.H. Coombs, *A Theory of Data* (2. Aufl. New York u.a. 1967), S. 280.

Skalierung läßt sich dementsprechend als eine Variante des Guttman-Modells mit einem freien Parameter formalisieren.

Nur am Rande sei vermerkt, daß wahrscheinlichkeitstheoretische Modellierungen dieser Art mit Theorien unscharfer Mengen eng verwandt sind, die heute in Expertensystemen und speziell in der Entwicklung von Verfahren "quellennaher" Datenverarbeitung für Historiker eine große Rolle spielen.¹⁵ Sucht man nach geistesgeschichtlichen Anknüpfungspunkten für das Wirklichkeitsbild, das probabilistischen Skalierungsverfahren zugrundeliegt, dann kann man bis zu den Kategorien des "Wesentlichen" und des "Zufälligen" oder von "Wirklichkeit" und "Erscheinung" in der idealistischen Philosophie zurückgehen.

2. Das Verfahren der Mokken-Skalierung

Bevor wir das Verfahren der Mokken-Skalierung im folgenden näher beschreiben, ist zu klären, was unter dem Begriff der "Skala" verstanden werden soll. Ganz allgemein ist eine Skala ein Meßinstrument zur Erfassung der Ausprägungen einer bestimmten Eigenschaft von Objekten. Der Meßvorgang als solcher besteht dabei in der Zuordnung von Skalenwerten zu den Objekten, wobei die Relation der Skalenwerte der Relation der Objekte entsprechen soll. Vielfach wird in diesem Zusammenhang die Welt der Objekte auch als "empirisches Relativ" bezeichnet, während man die Welt der Skalenwerte "numerisches Relativ" nennt.¹⁶

Der Begriff des "numerischen Relativ" ist jedoch insofern problematisch, als Zahlen nicht notwendig die einzige Möglichkeit zur Darstellung von Skalenwerten sind. Im Falle von "Nominalskalen" ist die Verwendung von Zahlenzeichen (Ziffern) zur Darstellung von Skalenwerten sogar irreführend - suggeriert sie doch Eigenschaften, die in diesem Fall den Skalenwerten nicht zukommen: Codiert man z.B. "politisch rechts" mit 1 und "politisch links" mit 2, so gilt zwar - wenn man 1 und 2 nicht als Ziffern, sondern als Zahlen auffaßt - daß 2 größer ist als 1, nur hat diese Relation kein empirisches Äquivalent, da eine Aussage "links" größer als "rechts" sinnlos ist. Noch absurder wäre eine Aussage wie: $2 \cdot \text{rechts} = 1 \cdot \text{links}$, obwohl $2 \cdot 1 = 1 \cdot 2$ arithmetisch sicherlich richtig ist. "Rechts" und "links" sind in diesem Zusammenhang lediglich zwei disjunkte Klassen von Objekten, die durch zwei beliebige verschiedene Sym-

¹⁵ Peter K. G. Günther, Die wahrscheinlichkeitstheoretischen Grundbestände der Testtheorie und verwandter formaler Gebilde, so auch der Latenten-Struktur-Theorie. Phil.Diss., Univ. Bremen, FB Psychologie, in Vorb.

¹⁶ Bernhard Orth, Einführung in die Theorie des Messens, 1974.

bole repräsentiert werden können. Die Verwendung zweier Ziffern ist hier nur eine von vielen Möglichkeiten - und mit Sicherheit eine problematische.

Bei der Messung von z.B. Längen oder Gewichten ist allerdings die Verwendung von Zahlen (dargestellt durch Ziffern) geradezu geboten, da sie die Eigenschaften des empirischen Relativs sehr gut abbilden können. Da hier sogar Aussagen wie "doppelt so lang" oder "halb so schwer" getroffen und repräsentiert werden können, spricht man in diesen Fällen von einer Verhältnis- oder Ratioskala. Wesentliches Merkmal von Ratioskalen ist, daß der Zahl 0 jegliches Fehlen der betreffenden Eigenschaft (z.B. Länge) zugeordnet ist. Hierin unterscheidet sich die Ratioskala von der Intervallskala. Für letztere gilt, daß zwar gleiche Abstände (Differenzen) zwischen zwei Skalenwerten (z.B. $3 - 1 = 4 - 2$) gleich große Unterschiede im empirischen Relativ repräsentieren, Verhältnisaussagen wie "doppelt so..." oder "halb so..." hingegen nicht aus den Skalenwerten abgeleitet werden können.

Ein Beispiel für eine Intervallskala ist die Temperaturskala nach Celsius. Eine Aussage wie: "Heute ist es doppelt so warm wie gestern" aufgrund der Messung von 4°C und 2°C ist nicht haltbar, da dies impliziert, daß bei 0°C keine Temperatur vorläge. "Keine Temperatur" aber ist - wie Chemie und Physik lehren - gleichbedeutend mit "keine Molekularbewegung" und dies entspricht etwa -273°C (oder 0°Kelvin - womit die Kelvin-Skala zu einer Ratioskala wird).

Sehen wir uns den Vorgang der Messung am Beispiel der Temperaturbestimmung mittels Thermometer etwas genauer an, so stellen wir fest, daß ein Thermometer im Grunde eine Längenskala aufweist, über die die Ausdehnung von Quecksilber erfaßt wird und von der auf das Ausmaß der Molekularbewegung - also die Temperatur - rückgeschlossen wird. Manifest und damit tatsächlich meßbar ist also nicht die Temperatur, sondern lediglich die Ausdehnung von z. B. Quecksilber, von der auf das Ausmaß eines hierfür verantwortlichen "latenten" Phänomens der Molekularbewegung geschlossen wird. Dies ist eine Situation, die in den empirischen Sozialwissenschaften vertraut ist. Ein Beispiel ist der Gegenstand dieses Beitrags: die politische Orientierung von Abgeordneten in historischen Parlamenten. Diese läßt sich nicht direkt erfassen, sondern kann nur aus manifestem Abstimmungsverhalten erschlossen werden. Politische Orientierung wird dabei als latente, das manifeste Verhalten determinierende Größe, als ein hypothetisches Konstrukt angesehen. Erschwerend ist im gegebenen Fall, daß das manifeste Abstimmungsverhalten durchaus auch andere Determinanten haben kann, wie etwa taktische Kalküle. Das Verhalten bei nur einer einzigen Abstimmung als Indikator für die politische Orientierung des jeweils Abstimmenden zu be-

trachten wäre deshalb genauso gefährlich, wie aus dem Verhalten mehrerer Personen bei nur einer Wahl auf die Determinanten des Wahlverhaltens schlechthin zu schließen. Zur Untersuchung der Determinanten des Abstimmungsverhaltens ist es also notwendig, mehr als eine Abstimmung einer Reihe von Abgeordneten heranzuziehen. Aber: Ist eine solche Sammlung von Abstimmungen *eo ipso* eine Skala?

Gibt man sich damit zufrieden, die Voten einer jeden Person in einem Antwortmustervektor zu protokollieren, diese untereinander auf Gleichheit zu prüfen und dabei lediglich "Vektoren gleich" bzw. "Vektoren ungleich" festzuhalten, so bildet die Menge der möglichen Antwortmustervektoren sicherlich eine (sehr umfangreiche) Nominalskala, nur ist der aus deren Anwendung zu erwartende Erkenntnisgewinn eher gering.

Will man nachweisen, daß politische Orientierungen wesentliche Determinanten des Abstimmungsverhaltens waren, ist es notwendig, zunächst aus der Gesamtmenge der Indikatoren (Abstimmungen) jene zu einem "Metaindikator" zusammenzufassen, die das hypothetische Konstrukt erfassen. Die "Metaindikatoren" sollen dabei den Charakter einer Dimension haben, entlang derer sich die konstituierenden Einzelindikatoren in einer eindeutigen Reihenfolge anordnen lassen. Gemeinsame Eigenschaften der Einzelindikatoren, hier also der Abstimmungsvorlagen, sollen dann zur inhaltlichen Identifikation der Dimension(en) herangezogen werden. Ferner soll die Position, die eine Person auf einer Dimension inne hat, eindeutig aus ihrem Antwortmustervektor bestimmt werden können, wobei dieser aus den Voten zu den jeweils konstituierenden Einzelindikatoren nach einer eindeutigen Vorschrift zusammengesetzt wird.

Diese Beschreibung der Aufgabe legt es nahe, hauptkomponentenanalytische Ansätze - wie die Faktorenanalyse - oder Verfahren aus dem Repertoire der Multidimensionalen Skalierung (MDS) heranzuziehen. Die Hauptkomponentenanalyse erfordert aber - da in ihr Datenwerte additiv verknüpft werden - Eingangsdaten von zumindest Intervallskalen-Niveau, und es wird - da dieses Verfahren lediglich die Parameter Varianz und Mittelwert berücksichtigt - Normalverteilung der Daten vorausgesetzt. Die vorliegenden Daten aber sind allenfalls ordinal (Ablehnung/Enthaltung/Zustimmung) - d.h. die Datenwerte einer jeden Abstimmung können zwar eindeutig je einer von drei Kategorien zugeordnet werden, welche ihrerseits in eine natürliche Ordnung zu bringen sind (sofern "Enthaltung" als eindeutiger Indikator für Indifferenz angesehen werden kann), eine Aussage über die Größe der Abstände

"Ablehnung - Enthaltung", "Enthaltung - Zustimmung" und "Ablehnung - Zustimmung" ist jedoch nicht möglich. Von daher verbietet sich streng genommen jedwede additive Verknüpfung von Datenwerten und damit die Berechnung von Mittelwerten und Varianzen. Zur Illustration: Eine Codierung Ablehnung = - 1, Enthaltung = 0, Zustimmung = + 1 bildet die uns bekannten Gegebenheiten im empirischen Relativ genauso gut oder schlecht ab wie eine Codierung Ablehnung = - 1, Enthaltung = 5, Zustimmung = 10 000.

Man könnte also auch sagen, unser Meßinstrument ist für den Einsatz hauptkomponentenanalytischer Verfahren zu ungenau. Hier liegt dann aber auch der Ansatzpunkt für eine heuristische Begründung, nach der diese Verfahren doch eingesetzt werden können: Codiert man nämlich Enthaltung = x , Ablehnung = $x-y$, Zustimmung = $x+y$ und setzt man voraus, daß Ablehnung ungefähr genauso weit von Enthaltung entfernt ist wie Zustimmung (Betrag) und betrachtet man die durch die Grobheit des Meßinstrumentes verursachten Fehler in den Einzelwerten als normalverteilte Zufallsgröße, so kann man den Einsatz von z.B. hauptkomponentenanalytischen Verfahren rechtfertigen.

Die Multidimensionale Skalierung wiederum wirft insbesondere das Problem auf, daß hier eine sog. Stressfunktion minimiert werden muß, die u.U. neben einem globalen Minimum etliche lokale Minima aufweisen kann, in welchen sich der Minimierungsalgorithmus "verfangen" kann, ohne daß dies feststellbar wäre.

Benötigt wird aufgrund der vorliegenden Datenqualität ein Verfahren, das es erlaubt, aus ordinalen Eingangsdaten Metaindikatoren zu bilden, ohne dabei Annahmen über die Meßeigenschaften der Daten zu überfordern oder Probleme verfahrenstechnischer Imponderabilien aufzuwerfen. Ein solches Verfahren ist die Skalenanalyse nach Mokken,¹⁷ die wiederum aufgefaßt werden kann als eine probabilistische Verallgemeinerung der Skalogrammanalyse nach Guttman.¹⁸ Im folgenden wollen wir zunächst kurz auf die Guttman-Skalogrammanalyse für dichotome Daten eingehen, von dieser dann auf die Mokken-Skalierung dichotomer Daten übergehen und diese schließlich auf polytome Daten ausweiten.

Nehmen wir also zunächst an, wir hätten es mit dichotomen Voten zu tun - es gebe also nur die Kategorien "Ablehnung = 0" und "Zustimmung = 1". Gäbe es nun nur eine einzige das Abstimmungsverhalten determinierende

¹⁷ R.J. Mokken, A Theory and Procedure of Scale Analysis (Den Haag 1971).

¹⁸ Vgl. Anm. 13.

latente Dimension, so müßten sich die Gesetzesvorlagen genauso wie die Abstimmenden entlang dieser Dimension in eine eindeutige Reihenfolge bringen lassen. Dies könnte wie folgt geschehen: Zunächst wird die Matrix aller aufgetretenen disjunkten Antwortmustervektoren aufgestellt, wobei diese Vektoren die Zeilen der Matrix bilden, so daß für jede Gesetzesvorlage genau eine Spalte existiert (vgl. Abb. 1a).

Abbildung 1a: Hypothetisches Abstimmungsmuster

| | | Vorlagen | | | | | |
|----------------------------|---|----------|---|---|---|-------|----------------------------------|
| | | 1 | 2 | 3 | 4 | 5 ... | |
| M u s t e r | 1 | 0 | 1 | 1 | 1 | 0 ... | (1: Zustimmung, 0: Ablehnung) |
| | 2 | 0 | 1 | 0 | 1 | 0 ... | |
| | 3 | 0 | 0 | 0 | 1 | 0 ... | |
| | 4 | 1 | 1 | 1 | 1 | 0 ... | |
| | 5 | 1 | 1 | 1 | 1 | 1 ... | |
| | | | | | | | |
| | | | | | | | |

Kann nun eine Zeilen- und Spaltenpermutation wie in Abb. 1b dargestellt gefunden werden, so liegt eine perfekte Guttman-Skala vor. Die Skalenposition einer Gesetzesvorlage entspricht ihrer neuen Spaltennummer, die eines Abstimmenden der neuen Zeilennummer seines Antwortmustervektors in der permutierten Matrix.

Abbildung 1b: Hypothetisches Abstimmungsmuster bei perfekter Guttman-Skala

| | | Vorlagen | | | | | |
|----------------------------|-------|----------|-----|-----|-----|-----|---|
| | | 1 | 2 | 3 | 4 | 5 | |
| | | (4) | (2) | (3) | (1) | (5) | In Klammern: Zeilen- und Spaltennummern von Tab. 1a |
| M u s t e r | 1 (3) | 1 | 0 | 0 | 0 | 0 | |
| | 2 (2) | 1 | 1 | 0 | 0 | 0 | |
| | 3 (1) | 1 | 1 | 1 | 0 | 0 | |
| | 4 (4) | 1 | 1 | 1 | 1 | 0 | |
| | 5 (5) | 1 | 1 | 1 | 1 | 1 | |
| | | | | | | | |
| | | | | | | | |

Die Praxis zeigt jedoch, daß eine perfekte Guttman-Skala eher die Ausnahme ist - d.h. selbst die beste aller Permutationen resultiert in einer Matrix, die Abweichungen von obigem Schema enthält. Die Skala ist dann nicht perfekt. Definiert man z.B.

$$\text{REP} = \frac{E}{N * I} \quad (\text{Gleichung 1})$$

mit: E - Anzahl der Fehler (1 statt 0, 0 statt 1)
 N - Anzahl der Antwortmustervektoren (Abstimmende)
 I - Anzahl der Gesetzesvorlagen

so gelangt man zu einem möglichen Maß für die Güte einer Skala. Dieses sollte - einem Vorschlag von Guttman folgend - nicht kleiner als 0.90 sein (entspr. max. 10% Fehler über alle Vorlagen und Abstimmenden hinweg). Inzwischen wurden neben dem REP-Koeffizienten weitere Gütemaße für Guttman-Skalen entwickelt, ohne daß dies jedoch die Logik des Verfahrens verändert.

Die Skalogrammanalyse nach Guttman kann also wie folgt zusammengefaßt werden:

1. Bilde alle möglichen Spalten- und Zeilen- Permutationen der Matrix der disjunkten Antwortmustervektoren und wähle diejenige aus, für die der REP-Koeffizient maximal ist.
2. Prüfe, ob REP größer als 0. 90 ist.
 Wenn ja: Fertig
 Wenn nein: Eliminiere eine Spalte und gehe zu 1.

Auch wenn über die Einführung beispielsweise des REP-Koeffizienten ein gewisses Maß an Fehlern zugelassen wird, geht dieser Ansatz von Guttman doch davon aus, daß die latente Eigenschaft das manifeste Verhalten eindeutig determiniert. Damit ist dieser Ansatz den deterministischen Verfahren zuzuordnen. Dieser Klasse stehen die probabilistischen Verfahren gegenüber, denen ein Denkmodell zugrunde liegt, demzufolge mit steigendem Ausmaß der latenten Eigenschaft die Wahrscheinlichkeit für ein bestimmtes manifestes Verhalten monoton steigt. Diese Wahrscheinlichkeit kann sich beliebig nahe an 1 heranbewegen, erreicht 1 jedoch nie ganz, so daß immer eine modellimmanente Möglichkeit für Abweichungen besteht, wiewohl diese Möglichkeit beliebig klein sein kann.

Als probabilistische Form der Guttman Skalogrammanalyse wird die Skalenanalyse nach Mokken angesehen, die umfassend erstmals 1971 vorgestellt

wurde. Einzug nimmt die Probabilistik hier durch die Verwendung des H-Koeffizienten als Maß für die Güte einer Skala. Der Algorithmus des Verfahrens kann wie folgt umrissen werden:¹⁹

- 1.: Suche die beiden "homogensten" Abstimmungen der (Rest-) Gesamtheit aller (noch nicht skalierten) Abstimmungen.
 - 1a: Prüfe, ob für diese überhaupt von "Homogenität" gesprochen werden kann.
 Wenn nein: Keine (weitere) Skalenbildung möglich. Fertig.
 Wenn ja: Betrachte die beiden Abstimmungen als "Anfang" einer Skala und
- 2.: Versuche, die Skala um eine Abstimmung zu erweitern.
 - 2a: Schließe alle Abstimmungen für diese Skala endgültig aus, die einen negativen "Homogenitätskoeffizienten" mit der soweit gebildeten Skala aufweisen.
 - 2b: Prüfe, ob es Items gibt, die zur bisherigen Skala "homogen" sind.
 Wenn nein: Keine Erweiterung mehr möglich.
 Skala fertig. Gehe zu 4.
 Wenn ja: Weiter bei 3.
- 3.: Wähle von den zur bisherigen Skala "homogenen" Abstimmungen diejenige aus, die zur "homogensten" erweiterten Skala führt. Gehe zu 2.
- 4.: Versuche, aus den noch nicht skalierten Abstimmungen eine weitere Skala zu bilden. Gehe wieder zu 1.

In dieser Beschreibung des Skalierungsalgorithmus' nach Mokken tritt wiederholt der Begriff der Homogenität auf. Diese wird definiert und quantifiziert durch Loevingers H-Koeffizienten,²⁰ der für zwei dichotome Variablen (Abstimmungen) i und j wie folgt dargestellt werden kann: Seien die beiden Abstimmungen kodiert als: 0 = Ablehnung, 1 = Zustimmung und werde der Vorlage i seltener zugestimmt als der Vorlage j. Dann lassen sich die zugehörigen (kombinierten) Zustimmungswahrscheinlichkeiten in der in Abb. 2 dargestellten Form in einer Vier-Felder-Tafel anordnen.

¹⁹ Vgl. Anm. 17.

²⁰ J. Loewinger, A systematic approach to the construction and evaluation of tests of ability. Psychological monographs, 61 (4).

Abbildung 2: Kombinierte Zustimmungswahrscheinlichkeiten zweier Abstimmungen

Vorlage j

| | | 0 | 1 | |
|--------------|---|----------|----------|----------|
| Vorlage i | 0 | p_{i0} | p_{i1} | $p_{i.}$ |
| | 1 | p_{1i} | p_{1i} | $p_{1.}$ |
| | | $p_{.0}$ | $p_{.1}$ | $p_{..}$ |

mit: $p_{1.} \leq p_{i.}$

H_{ij} ist dann definiert als

$$H_{ij} = 1 - \frac{p_{i0}}{p_{1.} * p_{.1}} \quad (\text{Gleichung 2})$$

oder - bei Schätzung der p_{ij} aus den Zellenhäufigkeiten n_{ij} der Stichprobe

$$H_{ij} = 1 - \frac{n_{i0}}{(n_{1.} * n_{.1}) / n_{..}} \quad (n_{..} = N) \quad (\text{Gleichung 2'})$$

Nebenbei bemerkt entspricht dieser Koeffizient sowohl

$$PHI_{ierr} = \frac{PHI}{PHI_{iii}} \quad (\text{Gleichung 2''})$$

als auch dem DEL -Koeffizienten nach Hildebrand et al. mit (1;0) als Fehlerzelle²¹. Bei der Quantifizierung von Homogenität ist es nahezu zwingend, (1;0) als Fehlerzelle zu betrachten, da unter der Voraussetzung $p_{1.} \leq p_{i.}$ diese Zelle all jene Abstimmenden enthält, die der Vorlage mit der geringeren Annahmewahrscheinlichkeit zugestimmt haben, obwohl sie der Vorlage mit der global höheren Annahmewahrscheinlichkeit ihre Zustimmung entzogen haben - eine Konstellation, die bei eindeutiger Anordnung von Vorlagen und Abstimmenden entlang einer einzigen Dimension nicht auftreten dürfte. (Bilden zwei Vorlagen eine perfekte Guttman-Skala, so ist $n(1;0) = 0$ und H_{ij} wird 1).

²¹ Georg Rudinger, Friedrich M. Chaselon, Hans J. Henning, E. J. Zimmermann, Qualitative Methoden - Neue Wege sozialwissenschaftlicher Methodik, München 1985.

Anders betrachtet wird im H-Koeffizienten für zwei Items i und j, $p_{1.} \leq p_{.1}$, der Erwartungswert für die Fehlerwahrscheinlichkeit $p(1;0)$ in Beziehung gesetzt zu deren Erwartungswert unter der Nullhypothese (bei Vorhersage von $p(1;0)$ aufgrund der bloßen Kenntnis der Randverteilungen). Bezeichnet man diese mit E_{ij} bzw. mit EO_{ij} , so ist

$$H_{ij} = 1 - \frac{E_{ij}}{EO_{ij}} \quad (\text{Gleichung 3})$$

Hier wird deutlich, daß der H-Koeffizient 0 wird, wenn E_{ij} und EO_{ij} gleich sind, wenn also der Erwartungswert der Fehlerwahrscheinlichkeit unter der Homogenitätshypothese ($p(1;0)$) gleich deren Erwartungswert unter der Nullhypothese ($p(1;.) \cdot p(.;0)$) ist, so daß deren Quotient 1 ergibt. Es bleibt jedoch die Frage, inwieweit sich ein H-Koeffizient > 0 statistisch signifikant von 0 unterscheidet. Dies läßt sich auf die Frage reduzieren, inwieweit sich

$$E_{ij} - EO_{ij} = p_{10} - p_{1.} \cdot p_{.0}$$

signifikant von 0 unterscheidet. Hierzu konnte Mokken zeigen, daß folgende Teststatistik sich für steigende Stichprobengröße $N = n..$ einer Normalverteilung nähert:²²

$$\Delta^* = \frac{(p_{10} - p_{1.} \cdot p_{.0}) - 0}{S^1(p_{10})} = \frac{p_{10} - p_{1.} \cdot p_{.0}}{S^1(p_{10})} \quad (\text{Gleichung 4})$$

Bei Schätzung der Parameter aus einer Stichprobe ergibt sich hieraus:

$$\Delta^* = \hat{\Delta}^* = \frac{n_{10} - (n_{1.} \cdot n_{.0}) / N}{S^1(n_{10})} \quad (\text{Gleichung 4'})$$

Gleichung 2' und Gleichung 4' liefern uns alle Hilfsmittel zur Durchführung von Schritt 1 und Schritt 1a des o. a. Ablaufschemas der MOKKEN-Skalenanalyse:

- Die beiden homogensten Vorlagen sind diejenigen mit dem höchsten (positiven) H_{ij} -Koeffizienten (1.)

²² Vgl. Anm. 17.

- Dieser kann auf statistische Bedeutsamkeit (signifikante Abweichung von 0) geprüft werden (1a.)²³

Für die übrigen Schritte fehlen jedoch noch

- die Definition der Homogenität einer Abstimmung mit einer (Teil)-Skala (H_j)
- die Definition der Homogenität einer ganzen Skala (H) nebst deren Signifikanzprüfungen.

H wird definiert als

$$H = 1 - \frac{E_{..}}{E0_{..}} \quad (\text{Gleichung 5})$$

- Mit: $E_{..}$ - Summe der Fehlerwahrscheinlichkeiten $p(1;0)$ aller H_{ij} , $i < j$ der nach Zustimmungswahrscheinlichkeiten absteigend geordneten Vorlagen der Skala.
- $E0_{..}$ - Summe der zugehörigen Erwartungswerte unter der Nullhypothese.

H_j wird analog definiert zu:

$$H_j = 1 - \frac{E_{.j}}{E0_{.j}} \quad (\text{Gleichung 6})$$

- Mit: $E_{.j}$ - Summe der $p(1;0)$ aller H_{ij} , $i \neq j$
- $E0_{.j}$ - Summe der zugehörigen Erwartungswerte unter der Nullhypothese.

Auch für diese lassen sich approximativ normalverteilte Teststatistiken DELTA* ableiten, die einen Test auf signifikante Abweichung von 0 erlauben. In diesem Zusammenhang muß jedoch betont werden, daß eine signifikante Abweichung von Null lediglich besagt, daß die Wahrscheinlichkeit dafür, daß keine Homogenität vorliegt, sehr gering ist. Dies ist jedoch keinesfalls gleichbedeutend mit dem Vorliegen einer brauchbaren Skala. Mokken schlägt daher vor, a priori einen Mindest-Skalierbarkeitskoeffizienten für die zu findenden Skalen festzulegen.

²³ Da etliche simultane Tests durchgeführt werden, ist eine entsprechende Adjustierung des α -Niveaus (etwa durch $\alpha^* = \alpha/N$ Tests) erforderlich.

Hierzu klassifiziert er:

| | |
|----------------------|----------------------|
| $H < 0.30$ | "keine Skala" |
| $0.30 \leq H < 0.40$ | "schwache Skala" |
| $0.40 \leq H < 0.50$ | "mittelmäßige Skala" |
| $0.50 \leq H$ | "starke Skala" |

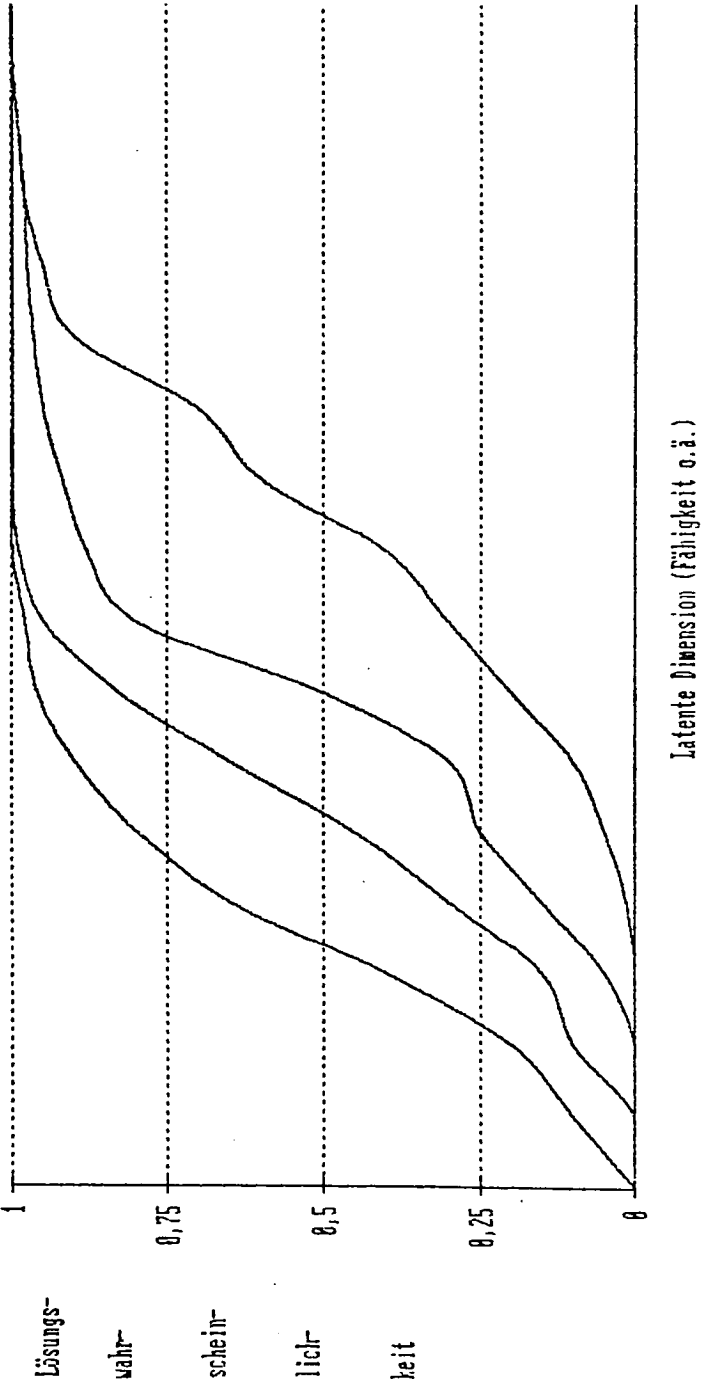
Allerdings wäre nach dem bislang Dargestellten eine solche Skala ungeachtet des H-Koeffizienten im Grunde nur ein Cluster homogener Variablen und nichts weiter. Um von einem Cluster zu einer Skala zu kommen, muß eine Ordnung in die konstituierenden Variablen - hier also in die Abstimmungen - gebracht werden, wobei sich als Ordnungskriterium das Komplement der Wahrscheinlichkeit für die 1-Kategorie - hier also der Zustimmungswahrscheinlichkeit für die Vorlagen - anbietet. Dabei gehen implizit folgende Annahmen ein:

1. Alle konstituierenden Variablen des Clusters erfassen die gleiche Eigenschaft.
2. Trägt man in einem Diagramm die Wahrscheinlichkeiten für die 1-Kategorien der konstituierenden Variablen gegen das Ausmaß der erfaßten Eigenschaft ab, so überschneiden sich die Funktionsgraphen der einzelnen Variablen (deren sog. "trace lines") nicht. D.h.: Für jede gegebene Wahrscheinlichkeit außer 0 und 1 der 1-Kategorie (hier: Zustimmungswahrscheinlichkeit) ist die Rangreihe der konstituierenden Variablen bezüglich des Ausmaßes der gemessenen Eigenschaft die gleiche (vgl. Abb. 3).

Sind diese beiden Annahmen haltbar und ordnet man die konstituierenden Variablen aufsteigend nach dem Komplement der Wahrscheinlichkeit der 1-Kategorie ($1-p_1$) bzw. absteigend nach p_1 , so ist der Skalenwert eines Untersuchungsobjektes identisch mit der Position der letzten 1 in seinem solchermaßen geordneten Antwortmustervektor.

Eine zweite Möglichkeit der Ermittlung des Skalenwertes eines Untersuchungsobjektes besteht darin, die Anzahl der Einsen im jeweiligen Antwortmustervektor auszuzählen. Im Idealfall - nämlich dann, wenn der geordnete Antwortmustervektor die Form $(1, \dots, 1, 0, \dots, 0)$ hat - sind die Ergebnisse beider Verfahren identisch. Bei Abweichungen vom Idealfall werden im ersten Verfahren die Skalenwerte eher überschätzt, während das zweite Verfahren eher unterschätzte Werte liefert.

Abbildung 3



Überträgt man diese beiden Vorschriften zur Ermittlung von Skalenwerten auf die vorliegenden Abstimmungsdaten, so können sie in entsprechender Terminologie wie folgt formuliert werden:

- a) Der Skalenwert eines Abgeordneten entspricht der Position der letzten Eins in seinem nach Zustimmungswahrscheinlichkeit der konstituierenden Vorlagen absteigend geordneten Abstimmungsmustervektor.
- b) Der Skalenwert eines Abgeordneten entspricht der Anzahl der Zustimmungen zu den konstituierenden Vorlagen.

Beide Verfahren liefern Skalenwerte auf ordinalem Meßniveau. Im Fall vieler fehlender Datenwerte liefern beide Verfahren in toto stark überschätzte Skalenwerte. Hier empfiehlt es sich, das zweite Verfahren dahingehend zu modifizieren, anstelle der einfachen Auszählung der Einsen deren relative Häufigkeit bezogen auf die Gesamtzahl der jeweils in einem Antwortvektor tatsächlich vorhandenen Daten zu benutzen.

In jedem Fall aber sollten die beiden o.a. impliziten Annahmen geprüft werden. Durch das Verfahren der Mokken-Skalierung ist automatisch gewährleistet, daß alle konstituierenden Variablen einer Skala die gleiche Eigenschaft oder ein "homogenes Korrelat" erfassen. Eine Unterscheidung von Eigenschaft und Korrelaten ist verfahrenstechnisch nicht möglich und kann daher nur inhaltlich-interpretativ erfolgen. Es empfiehlt sich daher, von vorneherein nur solche Variablen zu berücksichtigen, die unter einem inhaltlichen Gesichtspunkt gesehen vermutlich die gleichen latenten Dimensionen zu erfassen erlauben.

Die zweite der beiden o.a. impliziten Annahmen wird als doppelte Monotonie der trace lines (monoton steigend und sich nicht schneidend) bezeichnet. Zur Prüfung dieser Annahme genügt es nach Mokken,²⁴ die sog. P^{++} - und P^{--} -Matrices zu inspizieren. Bei der P^{++} -Matrix handelt es sich um die Matrix der p_{11} -Wahrscheinlichkeiten aller konstituierenden Variablen einer Skala. Die p_{11} -Wahrscheinlichkeit zweier Variablen i und j ist die Wahrscheinlichkeit für ein (1; 1) - Datenwertepaar in diesen beiden Variablen - in unserem Falle also die Wahrscheinlichkeit, in zwei Abstimmungen i und j gleichermaßen mit "ja" zu stimmen. Die P^{--} -Matrix ist analog die Matrix der Wahrscheinlichkeiten für (0;0)- Datenwertepaare. Ist nämlich doppelte Monotonie der trace lines gegeben und ordnet man Spalten und Zeilen dieser

²⁴ Vgl. Anm. 17.

beiden Matrices entsprechend der Wahrscheinlichkeiten für 1-Kodierungen in den Zeilen- und Spaltenvariablen absteigend an, so muß sich in der P^{++} -Matrix ein spalten- und zeilenweise monoton fallender Trend zeigen, während in der P^{-} -Matrix ein spalten- und zeilenweise monoton steigender Trend auftreten muß. Trendverletzungshäufigkeiten in einer Größenordnung von 3% und kleiner betrachtet Mokken dabei als irrelevant.²⁵

Ein wesentlicher Nachteil der Mokken-Analyse in der bislang dargestellten Form ist ihre Beschränkung auf dichotome Variablen. Diese Beschränkung wurde durch die Arbeiten von Molenaar aufgehoben.²⁶ Molenaar geht davon aus, daß mehrstufige (polytome) Variablen mit $M+1$ Kategorien als per se homogene "Mini-Skalen" bestehend aus M Schritten aufgefaßt werden können. Voraussetzung hierfür ist natürlich, daß die einzelnen Variablen mindestens ordinales Meßniveau aufweisen. Für K homogene Variablen mit jeweils M_k Schritten ergeben sich hieraus $\sum M_k$ Skalenschritte in der Gesamtskala - bildlich gesprochen durch "Verschachteln" oder "Ineinanderschieben/Aneinanderhängen" der Schritte der Einzelvariablen. Die Reihenfolge der Einzelvariablenschritte $S1_{1-3}$, $S2_{1-3}$ und $S3_{1-3}$ dreier dreistufiger Variablen könnten z. B. wie folgt entlang einer gemeinsamen latenten Dimension angeordnet sein:

$$S1_1 S2_1 S2_2 S1_2 S2_3 S3_1 S3_2 S3_3 S1_3$$

Allerdings sind die Schritte innerhalb einer einzigen Variablen keineswegs unabhängig voneinander. Würde man die Einzelvariablenschritte m und $m-1$ in zwei dichotome Dummy-Variablen codieren und diese in einer Vierfeldertafel gegenüberstellen, so könnte es nie vorkommen, daß die Zelle ($S_m = 1$; $S_{m-1} = 0$) eine Besetzung aufweist, da der ordinale Charakter der Variablen impliziert, daß immer dann, wenn Schritt m innerhalb einer Variablen vollzogen wurde (also die $(m+1)$ -te Stufe erreicht ist) auch alle vorangegangenen Schritte vollzogen wurden (also alle vorangegangenen Stufen bereits überschritten sind). Molenaar nennt diese Zellen daher folgerichtig "logisch unmögliche Fehlerzellen". Aus diesem Grunde aber ist es nicht möglich, eine Variable mit M_k Schritten einfach durch M_k dichotome Dummy-Variablen zu ersetzen und anschließend den o.a. Algorithmus für dichotome Variablen auf die Gesamtheit aller Dummy-Variablen anzusetzen. Betrachtet man dieses Problem vor dem Hintergrund der Arbeiten von Hildebrand et al. zum DEL-Koeffizienten, so reduziert sich dieses auf die Festlegung der Fehlerzel-

²⁵ Ebd., S. 277.

²⁶ Ivo W. Molenaar, Mokken scaling revisited, Kwantitative Methoden (3), 1982, S.145-164.

len in den bivariaten Kreuztabellen aller Variablenpaare sowie die Auswahl geeigneter Homogenitätsmaße.²⁷

Zur Festlegung der Fehlerzellen führt Molenaar das Konzept der "Cumulated Percentages Past" (CPP) ein, auf deren Basis in Übereinstimmung mit dem Guttman-Modell für mehrstufige Variablen die Sequenz der Wertepaare absteigender Auftretenswahrscheinlichkeit ermittelt wird. Unter den CPP sind dabei die kumulierten relativen Häufigkeiten der Randverteilungen der Kreuztabellen je zweier Variablen zu verstehen, wobei die Kumulation von der Endstufe mit der geringsten Auftretenswahrscheinlichkeit ausgehend erfolgt. Bei der oben beschriebenen Mokken-Analyse für dichotome Variablen wurde die Fehlerzelle immer a priori unabhängig von den Randverteilungen festgelegt und negative H_i hatten zur Folge, daß diese Variablen nicht in eine Skala aufgenommen wurden. Konsequenz dieser Vorgehensweise ist, daß lediglich unipolare Skalen generiert werden können. Bipolare Skalen treten daher immer in der Form von zwei unipolaren Skalen zutage - es sei denn, die konstituierenden Variablen eines der beiden Pole werden vorab umcodiert ($1 \rightarrow 0$, $0 \rightarrow 1$). Eine Orientierung an den relativen Häufigkeiten der Endstufen der Einzelvariablen resultiert jedoch in einer automatischen Umcodierung der konstituierenden Variablen eines evtl. negativen Poles - und zwar auch dann, wenn der Algorithmus für polytome Variablen auf dichotome Variablen angewendet wird.

Dieser Effekt ist zwar erwünscht, muß aber andererseits bei der Bestimmung der Skalenwerte einzelner Untersuchungsobjekte adäquat berücksichtigt werden. Die generelle Vorgehensweise der Ermittlung der Fehlerzellen im polytomen Fall läßt sich nur schwer allgemein darstellen, so daß sie anhand eines Beispielles illustriert werden soll. Abbildung 4 gibt die Kreuztabellen der relativen Häufigkeiten zweier Variablen i und j wieder. Betrachten wir zunächst die mit "Gesamt" bezeichneten Spalten- und Zeilensummen, so ist offensichtlich, daß für beide Variablen der Code 2 eine geringere Auftretenswahrscheinlichkeit hat als der Code 0. Die Ermittlung der CPP muß also durch Kumulation "von 2 nach 0" erfolgen - zeilenweise also von unten nach oben und spaltenweise von rechts nach links.

²⁷ D.K. Hildebrand, J.D. Laing & H. Rosenthal, Prediction Analysis of Cross Classifications, New York, Wiley, 1977.

Abbildung 4

| | | Variable i | | | Gesamt | CPP |
|----------|---|------------|----|----|--------|-----|
| | | 0 | 1 | 2 | | |
| Variable | 0 | 29 | 20 | 1 | 50 | 100 |
| j | 1 | 1 | 8 | 1 | 10 | 50 |
| | 2 | 1 | 12 | 27 | 40 | 40 |
| Gesamt | | 31 | 40 | 29 | 100 | |
| CPP | | 100 | 69 | 29 | | |

(Diese Kumulationsrichtungen indizieren, daß zur Erreichung der Stufe 2 der Variablen ein höheres Ausmaß der latenten Eigenschaft erforderlich ist, als zur Erreichung der Stufe 0.) Die erste erlaubte Zelle ist nun die Zelle mit der CPP-Kombination (100; 100). (Die theoretische Wahrscheinlichkeit dafür, daß alle Untersuchungsobjekte diese oder eine höhere Stufe der aus beiden Variablen gebildeten Skala erreicht haben, ist 1.0). Nun muß entschieden werden, welches die nächste erlaubte Zelle ist. Zur Auswahl stehen (0; 1) und (1;0). Die zu beantwortende Frage ist im Grunde: Ist der Schritt von 0 nach 1 kleiner für Variable i oder für Variable j ? Ein Blick auf die CPP zeigt, daß bei Variable i 69% der Untersuchungsobjekte die Stufe 1 oder eine höhere erreicht haben, während bei Variable j dies nur für 50% zutrifft. Der Schritt Von 0 nach 1 ist daher wohl bei Variable i kleiner als bei Variable j. Die nächste erlaubte Zelle ist damit die Zelle (0;1). Von dieser kann nun entweder der Schritt nach 2 in Variable i der nächstkürzere sein, oder der nach 1 in Variable j. Als nächste erlaubte Zelle stehen also zur Auswahl (0;2) und (1;1). Betrachten wir auch hier wieder die korrespondierenden CPP-Werte, so sind dies 29% für (0;2) und 50% für (1;1). Daraus wird geschlossen, daß der Schritt nach j = 1 der kürzere ist, so daß die Wahl auf (1;1) fällt. Nun stehen zur Wahl die Zellen (1;2) und (2;1) mit den korrespondierenden CPP-Werten 29% und 40%. Da $40 > 29$, fällt die Wahl auf (2;1), von wo aus nur noch (2;2) zur Auswahl steht.

Als erlaubte Zellen sind damit festgelegt: (0;0), (0;1), (1;1), (2;1) und (2;2). Die übrigen Zellen - also (0;2), (1;0), (1;2), (2;0) - liegen damit als Fehlerzellen fest und der Homogenitätskoeffizient H_{ij} kann durch Summation über diese Zellen hinweg (in Zähler und Nenner) analog zu Gleichung 3 bestimmt werden. In gleicher Analogie (zusätzliche Summation über alle Fehlerzellen hinweg) können die Koeffizienten H_i und $H_{..}$ bestimmt werden.

Der im Vorangegangenen vorgestellten Vorgehensweise zur Ermittlung der Fehlerzellen liegt die Annahme zugrunde, daß die beiden Variablen sich

entlang einer Dimension skalieren lassen. Ist diese Annahme nicht haltbar, so ist die Besetzung der Fehlerzellen verhältnismäßig hoch und der H-Koeffizient wird schlecht. Die Annahme der unidimensionalen Skalierbarkeit findet ihren Niederschlag in der Vorschrift, sich von einer Ecke der Kreuztabelle nur durch Einzelschritte in der Horizontalen oder in der Vertikalen in Richtung auf die diagonal gegenüberliegende Ecke bis hin zu dieser bewegen zu können.

Da es Molenaar desweiteren gelang, auch die Teststatistik DELTA* entsprechend zu verallgemeinern, kann der eigentliche Skalierungsalgorithmus unverändert vom dichotomen Fall übernommen werden.

3. Die Anwendung der Mokken-Skalierung auf die namentlichen Abstimmungen in der Frankfurter Nationalversammlung (FNV) und der Assemblée nationale constituante (ANC): Ergebnisse und Interpretation

Das Verfahren der Mokken-Skalierung wurde exemplarisch auf zwei Datenbestände angewendet, in denen das Verhalten von jeweils 900 (ANC) und 766 (FNV) Abgeordneten in 90 (FNV) und 86 (ANC) namentlichen Abstimmungen erfaßt wurde. Diese Voten sind systematische Zufallsauswahlen aus den 259 (ANC) und 298 (FNV) namentlichen und in den offiziellen Parlamentsberichten dokumentierten Abstimmungen. Erfaßt wurde - nach einem Zufallsstart - jeweils jede dritte Abstimmung in zeitlicher Folge. Anhand eines Datensatzes, der nahezu alle Abstimmungen in der FNV enthält, konnte nachträglich die Stichprobenvalidität überprüft werden. Es zeigte sich, daß das ungewöhnliche Verfahren einer Zufallsauswahl von Items keine erheblichen Verzerrungen bewirkte: Die Hauptskalen, die aus verschiedenen Stichproben und dem Datensatz mit (nahezu) allen Abstimmungen gewonnen wurden, korrelierten mit Werten von $r = 0.95$ und 0.99 .²⁸

Vor der Anwendung der Mokken-Skalierung wurden die Primärdaten in zweierlei Weise aufbereitet:

- (1) Zunächst wurden die Abstimmungen nach dem Kriterium der Zugehörigkeit (bzw. Affinität) der Antragsteller zu Fraktionen und Protofraktionen "gepolt". Entsprechende Informationen waren in den Parlamentsberichten und in den biographischen Materialien über die einzelnen Abgeordneten enthalten. Der Code +1 wurde der Ja-Stimme dann zugewiesen, wenn der Antrag von einem Abgeordneten der Rechten eingebracht wurde, der Code -1,

²⁸ Best, Bildung und Besitz, S. 498.

wenn ein Abgeordneter der Linken Antragsteller war. Bei Stimmenthaltung wurde Code 0 vergeben. In den folgenden Analysen werden die Daten in der Weise dichotomisiert, daß die Enthaltungen der Kategorie "Nichtzustimmung" zugewiesen wurden.²⁹

(2) Fehlende Werte, die auf Abwesenheit von Abgeordneten bei Abstimmungen zurückzuführen waren, wurden durch "Imputation" mittels multipler schrittweiser Regression geschätzt. Die Anwendung eines fall- und variablenorientierten Schätzverfahrens war geboten, da andere Methoden der Behandlung fehlender Werte, wie etwa deren pauschale Ersetzung durch einen "neutralen" Wert, zu schwerwiegenden Verzerrungen führen würden. So würden etwa Abgeordnete, die tatsächlich dem rechten oder linken Flügel angehört hatten, nach ihrem Abstimmungsverhalten der "Mitte" zugeordnet, wenn häufig bei Abwesenheit ein neutraler Wert vergeben worden wäre. Auch ergab bei der Schätzung fehlender Werte mit Regressionstechniken die Validitätsüberprüfung im Vergleich zu anderen missing-data-Techniken eine deutlich bessere Anpassung der Skalenwerte an die Kriteriumsvariable "Fraktionszugehörigkeit".³⁰

Die Analysen des Abstimmungsverhaltens mit dem Verfahren der Moken-Skalierung hatten ein prägnantes Ergebnis, das in den Grundzügen zwischen beiden Untersuchungsgruppen übereinstimmt. Wählt man den hohen Wert von $H > 0.50$ als Einschlußkriterium für die Aufnahme einer Abstimmung in eine Skala, dann lassen sich in der Frankfurter Nationalversammlung 79 von 90 (=88%) der Voten und in der Assemblée nationale constituante 66 von 86 Abstimmungen (=77%) jeweils einer Hauptskala zuordnen. Die für die Hauptskalen errechneten Gesamt-Homogenitätskoeffizienten erreichen sehr hohe Werte (ANC und FNV: $H = 0.76$), gleiches gilt für die Prüfstatistik DELTA* (ANC = 752.56; FNV = 946.91).

Es wurde jedoch schon darauf hingewiesen, daß die Homogenität eines Item-Satzes nicht hinreicht, um zu entscheiden, ob eine Skala vorliegt, die eine Eigenschaftsdimension abbildet. Dazu muß zusätzlich geklärt werden, ob die in die Hauptskala aufgenommenen Abstimmungen einem gemeinsamen "Bedeutungsuniversum" zugehören, das heißt, einem "Sinnzusammenhang, ... für den ein konkreter Satz von Indikatoren als operationale Definition verstanden werden kann".³¹

²⁹ Stimmenthaltungen waren in beiden Parlamenten außerordentlich selten und machten weniger als 0.3% der abgegebenen Voten aus. Ebd., S. 499.

³⁰ Heinrich Best, "Computing the Unmeasurable: Estimating Missing Values in Legislative Roll-Call Analysis", in: Peter Denley et al. (Hrsg.), *History and Computing*, Bd. 2 (Manchester u. New York 1989), S. 104-116.

In Abstimmungsanalysen ist es sonst üblich, diesen Sinnzusammenhang aus den manifesten Inhalten der Vorlagen zu erschließen. Das ist ein problematisches Vorgehen, weil sich dabei gewöhnlich ein hoher Arbeitsaufwand mit einem unerfreulich großen Ermessensspielraum bei der Deutung von Abstimmungstexten verbindet. Im gegebenen Fall verfügen wir jedoch über leistungsfähige externe Kriteriumsvariablen, die eine zuverlässigere Zuordnung ermöglichen: In beiden Versammlungen bildeten sich mehr oder weniger ephemere Fraktionen oder "Protofraktionen", die von den Beteiligten und den zeitgenössischen Beobachtern nach einem rechts-links-Schema geordnet wurden. Wenn die Skalen, wie erwartet, Positionen auf einem rechts-links-Kontinuum messen, dann müßten die Abgeordneten vom äußersten rechten Flügel über die Mitte zum linken Flügel hin abnehmende durchschnittliche Skalenwerte aufweisen. Die Skalenwerte der Abgeordneten wurden zuvor aus der Anzahl der positiven Voten gebildet, die für Vorlagen der Rechten abgegeben wurden. Wie erwähnt, ist dies ein konservatives Verfahren, weil es bei inkonsistentem Abstimmungsverhalten eher unterschätzte Werte liefert.³²

Das Ergebnis einer Varianzanalyse der Skalenwerte mit der Fraktionszugehörigkeit als Klassifikationsvariable unterstützt für beide Untersuchungsgruppen nachdrücklich die Annahme, daß die Skalen Positionen auf einem rechts-links-Kontinuum abbilden. Die durchschnittliche Anzahl positiver Reaktionen nimmt zu, je weiter rechts eine Fraktion oder eine Protofraktion plazierte war. Die historisch überlieferten und die aus dem Abstimmungsverhalten ermittelten Rangfolgen von Fraktionen und Protofraktionen sind nahezu identisch. Der in beiden Fällen sehr hohe ETA-Wert ($ANC = 0.80$; $FNV = 0.82$), der den Anteil der Summe der Abweichungsquadrate zwischen den Gruppen an den quadrierten Gesamtabweichungen angibt, zeigt an, daß die Gruppen relativ homogen sind. Wenn man - mit guten Gründen - jene Abgeordneten ausschließt, die keiner Fraktion angehörten oder für die die genauen Angaben zu den politischen Standorten im biographischen Material fehlen, verbessert sich dieser Wert erwartungsgemäß - im Fall der deutschen Untersuchungsgruppen sogar erheblich ($ANC = 0.84$; $FNV = 0.94$). In die gleiche Richtung weist die Inspektion der Standardabweichungen und der Quartilsabstände der Gruppenmittelwerte. Dabei fällt auf, daß die Streuung von den Flügelfraktionen zu den Mittelfraktionen hin zunimmt. Dies ist ein inhaltlich bedeutsamer Befund, der mit anderen Beobachtungen koinzidiert, nach denen die Fraktionsdisziplin der Abgeordneten bei den Mittelfraktionen geringer war.³³

³¹ Scheuch u. Zehnpfennig, Skalierungsverfahren, S. 117.

³² S.o., S. 60.

Alles in allem bestätigt aber die Korrelation der Skalen mit den Kriteriumsvariablen die Erwartung, daß ein links-rechts-Kontinuum in beiden Untersuchungsgruppen den gemeinsamen "Sinnzusammenhang" der Abstimmungen der Hauptskalen bildete. Die verbleibende Unschärfe ist - wie auch der Vergleich zwischen den Untersuchungsgruppen nahelegt - eher den Validitätsmängeln der Kriteriumsvariablen als denen der Skalen zuzuschreiben. Auch muß bedacht werden, daß die bei der Skalenbildung berücksichtigten Abstimmungen in einem langen Zeitraum stattfanden, während sich die Angaben über Gruppen- und Fraktionszugehörigkeiten auf bestimmte Zeitpunkte beziehen.

Eine Komplikation bei diesem Verfahren, das ja auf Mittelwertberechnungen beruht, besteht allerdings darin, daß Mokken-Skalen ebenso wie Guttman-Skalen lediglich die Meßeigenschaften von Ordinalskalen haben und demnach nur über Rangordnungen der Untersuchungseinheiten Auskunft geben. Wir haben uns jedoch aus drei Gründen berechtigt gefühlt, diese Beschränkung hier zu ignorieren. Zum einen werden unsere Skalen aus einer ungewöhnlich großen Anzahl von Items gebildet; die "ungemessenen" Abstände zwischen den Skalenwerten können also nicht allzu groß sein. Für diese Vermutung spricht auch, daß die Distanzen zwischen den Items, gemessen an ihrer "Schwierigkeit", nicht sehr groß sind. Zudem dient die Mittelwertberechnung hier ja lediglich der inhaltlichen Deutung und Validierung der Skalen, deren weitere Verwendung als metrische Indikatoren ist nicht beabsichtigt.

Aufschlüsse über die Mechanismen, die den Prozeß der Formierung von Konfliktgruppen steuerten, ergeben sich, wenn man die Besetzung der Skalenpositionen betrachtet (Abb. 5 u. 6). Dann zeigt sich ein markant bipolares Muster: In beiden Versammlungen konzentrierten sich die Fälle auf die äußeren Skalenpositionen. Dieses Ergebnis der Skalenanalysen unterstützt unsere Annahme, daß das Abstimmungsverhalten einer Entscheidungslogik folgte, bei der der status quo über Serien von Voten die gleichbleibende Alternative bildete. "Proximity-Voting" und strategisches Abstimmungsverhalten hätten dagegen zu überzufällig häufigen Inkonsistenzen in den Abstimmungsmustern oder einer stärkeren Besetzung des Mittelbereichs der Hauptskalen führen müssen. Die beobachteten Verteilungen sind dagegen ein starkes Argument für die These eines "natürlichen Bipartismus", wie sie u.a. Maurice Duverger vertreten hat: Das Ergebnis politischer Freiheit war nicht promiskuitive Pluralität, sondern ein rigoreses, wenn auch kompetitives, Kon-

³³ Best, Bildung und Besitz, S. 330 ff.

Abbildung 5

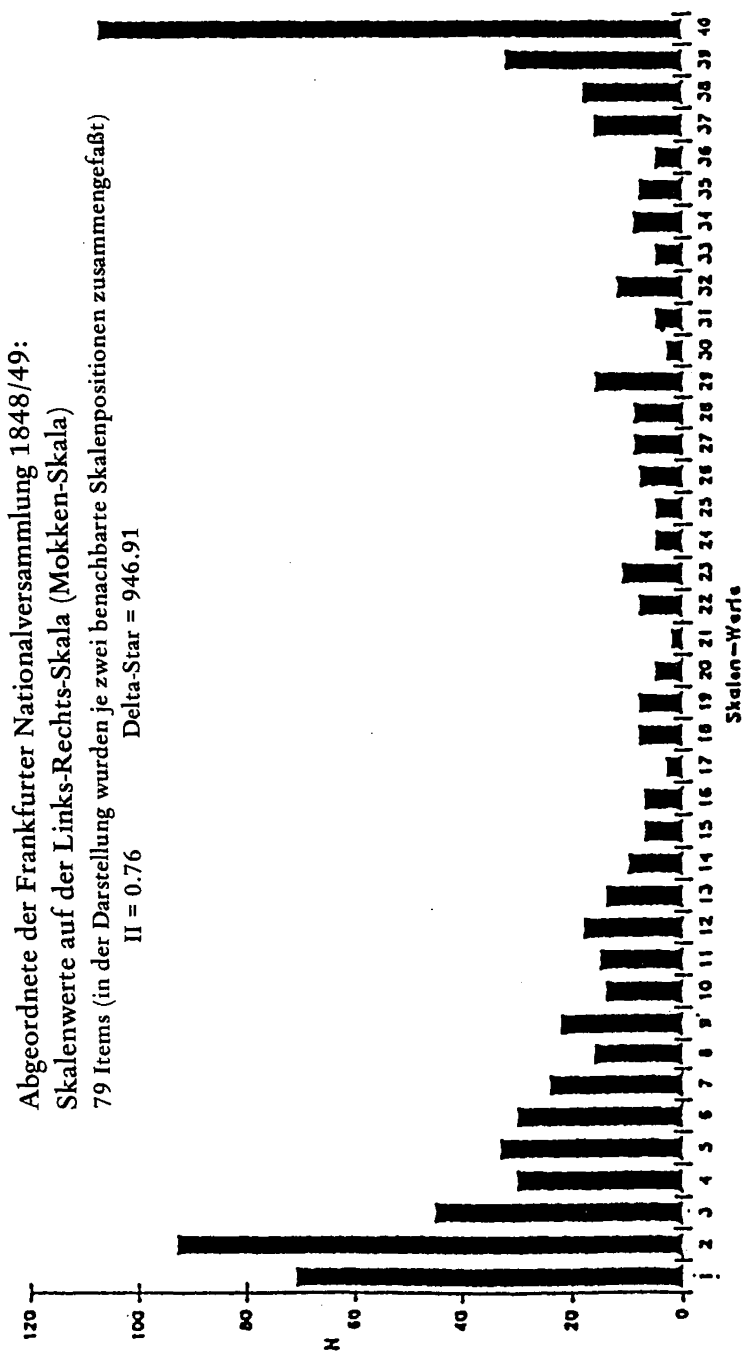


Abbildung 6

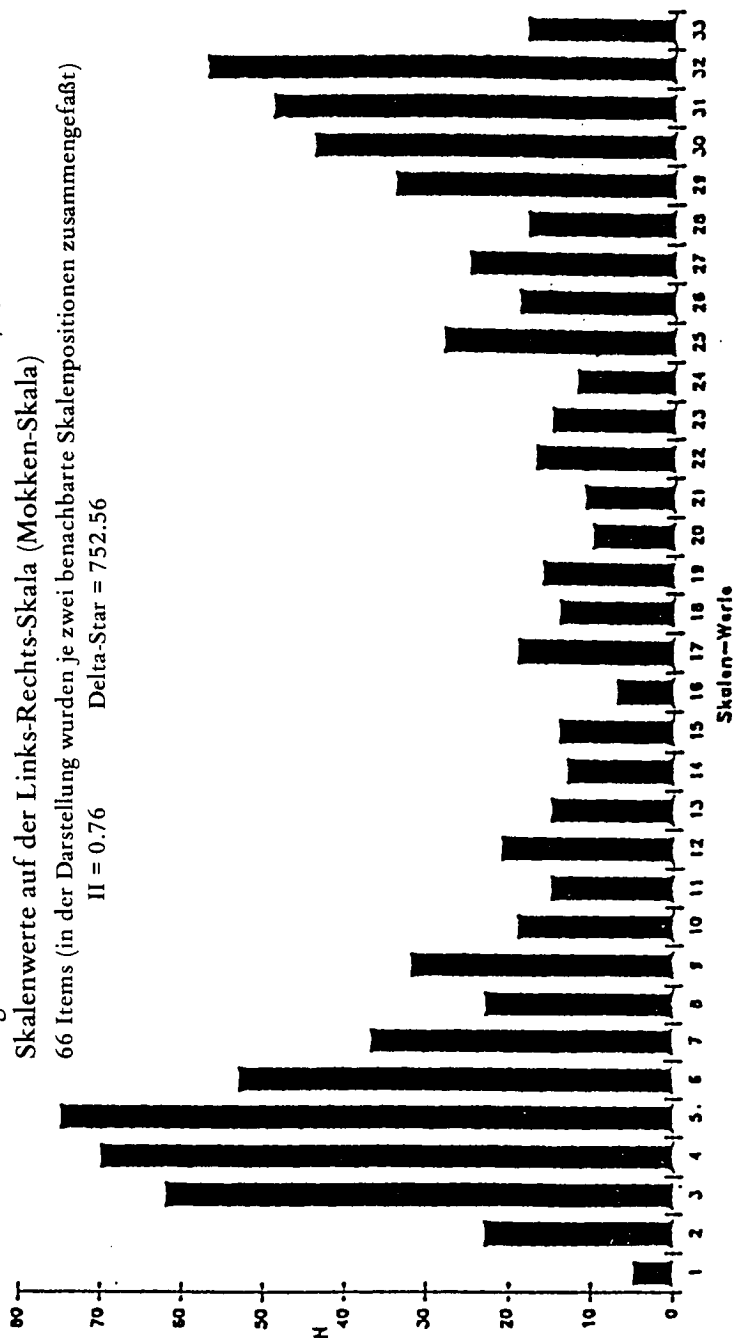
Abgeordnete der Assemblée Nationale Constituante 1848/49:

Skalenwerte auf der Links-Rechts-Skala (Mokken-Skala)

66 Items (in der Darstellung wurden je zwei benachbarte Skalenpositionen zusammengefaßt)

$\Pi = 0.76$

Delta-Star = 752.56



fliktmuster.³⁴

Die Frage stellt sich hier, warum in beiden Versammlungen, die in markant unterschiedlichen historisch-politischen Kontexten zusammentraten, isomorphe und rigide Konfliktstrukturen entstanden. Einen "Zwang zur Konsistenz", der in Befragungssituationen fehlt, setzte - wie erwähnt - sicherlich der Umstand, daß namentliche Abstimmungen öffentlich sind und Abweichungen der Abgeordneten von den Verhaltenserwartungen durch andere Akteure sanktioniert werden können. Doch erklärt dieser Sachverhalt nicht, warum sich in beiden Versammlungen bipolare Konfliktstrukturen ausprägten, was im übrigen ja auch den frühparlamentarischen Vorbehalten gegenüber "Parteiungen" jeder Art widersprach und die beobachteten Konstellationen zu einem Beispiel für ungeplante, ja unerwünschte Folge sozialen Handelns macht.

In einem anderen Zusammenhang wurde ein dynamisches Modell der Strukturbildung in parlamentarischen Versammlungen vorgeschlagen, das von der Überlegung ausgeht, daß die Formierung politischer Konfliktgruppen in Parlamenten den Koordinationserfordernissen kontinuierlicher Entscheidungsbildung unter kompetitiven Bedingungen folgt.³⁵ In vereinfachter Weise läßt es sich wie folgt darstellen: Ein primärer Prozeß der Gruppenbildung wird dadurch ausgelöst, daß Abstimmungen in Parlamenten - wie andere rationale Entscheidungen auch - Informations- und Koordinationskosten verursachen, die sehr hoch werden können, wenn sich Abgeordnete mit Serien von Vorlagen befassen müssen, zu deren Bewertung großes Fachwissen erforderlich ist oder die nur schwer absehbare politische Folgen haben. In solchen Situationen werden Abgeordnete versuchen, ihre Informationskosten zu senken, indem sie ihr Abstimmungsverhalten an solchen Kollegen orientieren, die sie für sachlich kompetent oder besonders vertrauenswürdig ansehen. Im Insider-Jargon der amerikanischen Parlamentarier wird dieser Vorgang mit "cue-taking" bezeichnet. Die Entstehung von Protofraktionen in Parlamenten läßt sich danach als das Ergebnis einer Spezialisierung und Arbeitsteilung deuten. (Proto-)Fraktionen bilden die Foren der Kommunikation, in denen Abgeordnete entweder ihr jeweiliges Expertenwissen austauschen oder Informationen über den Preis der Verpflichtung zu Konformität in Abstimmungen zu erlangen suchen.

³⁴ Les partis politiques. (Paris 1976 u.ö.), S. 293 ff. Auch Carl Schmitts Bestimmung des Politischen als ein polarisiertes Freund-Feind Verhältnis läßt sich hier als Referenz anführen.

³⁵ Ausführlich mit Literaturhinweisen: Best, Bildung und Besitz, S. 350-357.

Gelingt es, verschiedene Handlungsalternativen in einem niederdimensionalen Konfliktraum zu lokalisieren, führt dies zu einer weiteren kognitiven Entlastung der parlamentarischen Akteure. Robert Heiner hat darauf aufmerksam gemacht, daß unter turbulenten Handlungsbedingungen, wie sie ja in Revolutionen typischerweise bestehen, die Tendenzen zur Realitätsvereinfachung besonders ausgeprägt sind.³⁶

Solche Argumente begründen jedoch eher die Eindimensionalität der Konfliktstrukturen, weniger ihre Bipolarität, die wir vor allem dem Bemühen der Parlamentarier zurechnen, ihre Durchsetzungsmacht zu steigern. Die primären parlamentarischen Gruppen verstetigten sich und tendierten zur weiteren Aggregation infolge der parlamentarischen Mehrheitsregeln: stabile Mehrheiten über Zeit werden belohnt - das gilt selbst für die Verabschiedung von Einzelgesetzen mit zugehörigen Amendments.

Bedenkt man, daß politische Konflikte auch Kämpfe um Machtpositionen sind, wird eine weitere Knappheitsregel ins Spiel gebracht, die im Argument der "minimal winning coalition" formalisiert wurde:³⁷ Kleine Mehrheiten führen zu einer günstigeren Relation zwischen der Zahl der im Konkurrenzkampf um die Macht verfügbaren Positionen und der Zahl der durch Zugehörigkeit zur Mehrheitsgruppe zur Ämterappropriation berechtigten Aspiranten.

Schließlich gilt: eine bipolare Struktur erhöht die Mitgliederstabilität der parlamentarischen Konfliktgruppen, denn bei sonst gleichen Bedingungen ist in einem Zweiparteien-System für den Austritt oder das Ausscheiden im Durchschnitt ein höherer Preis zu zahlen als in einer Mehrparteien-Struktur.³⁸ Fassen wir die Argumente zusammen, dann erklärt sich die beobachtete eindimensional-bipolare Konfliktstruktur aus dem Zusammenwirken von drei Hauptfaktoren der Strukturbildung: (1) der Notwendigkeit, die Transaktionskosten bei der parlamentarischen Entscheidungsbildung zu senken, (2) der Tendenz, die Durchsetzungsmacht der parlamentarischen Konfliktgruppen zu steigern und dabei (3) einen möglichst starken Solidaritätsdruck auf die Gruppenmitglieder auszuüben.

Die Differenzierung und Verstetigung politischer Konfliktgruppen wird

³⁶ Ronald A. Heiner, "The Origin of Predictable Behavior", in: *American Economic Review*, 73 (1983), S. 560-589.

³⁷ L.C. Dodd, "Party Coalitions in Multiparty Parliaments. A Gametheoretic Analysis", in: *American Political Science Review*, 68 (1974), S. 1093-1117.

³⁸ Raymond Boudon, *Die Logik des gesellschaftlichen Handelns* (Neuwied u. Darmstadt 1980), S. 38; Michael Hechter, "A Theory of Group Solidarity", in: ders. (Hrsg.), *The Microfoundations of Macrosociology* (Philadelphia 1983), S. 16-57.

Abstimmungen in der Frankfurter Nationalversammlung
 Abbildung 7: H_j -Koeffizienten in sequentieller Ordnung (durchgezogene Linie:
 gleitende Mittelwerte, gestrichelte Linien: gleitende Standardabweichungen)

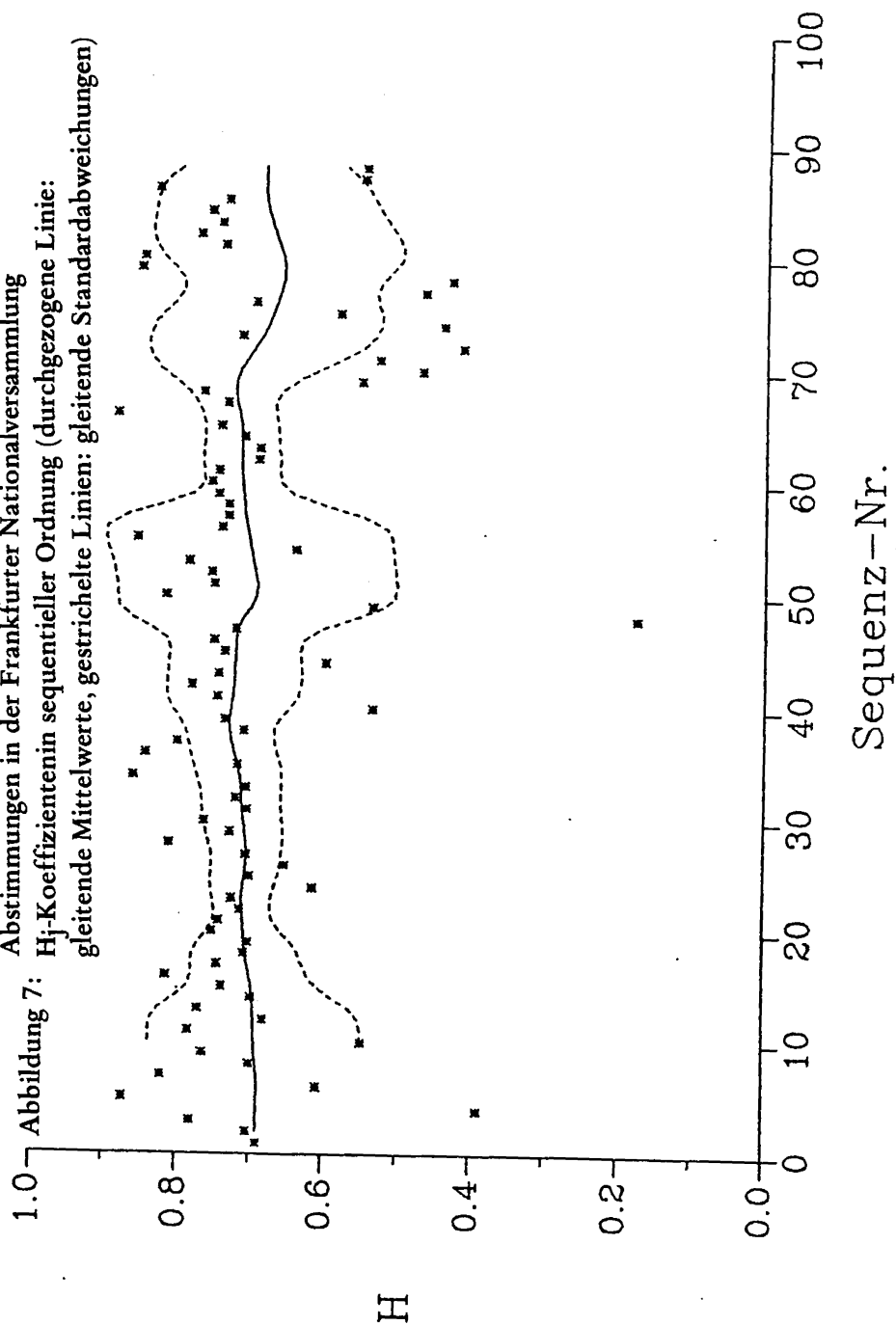
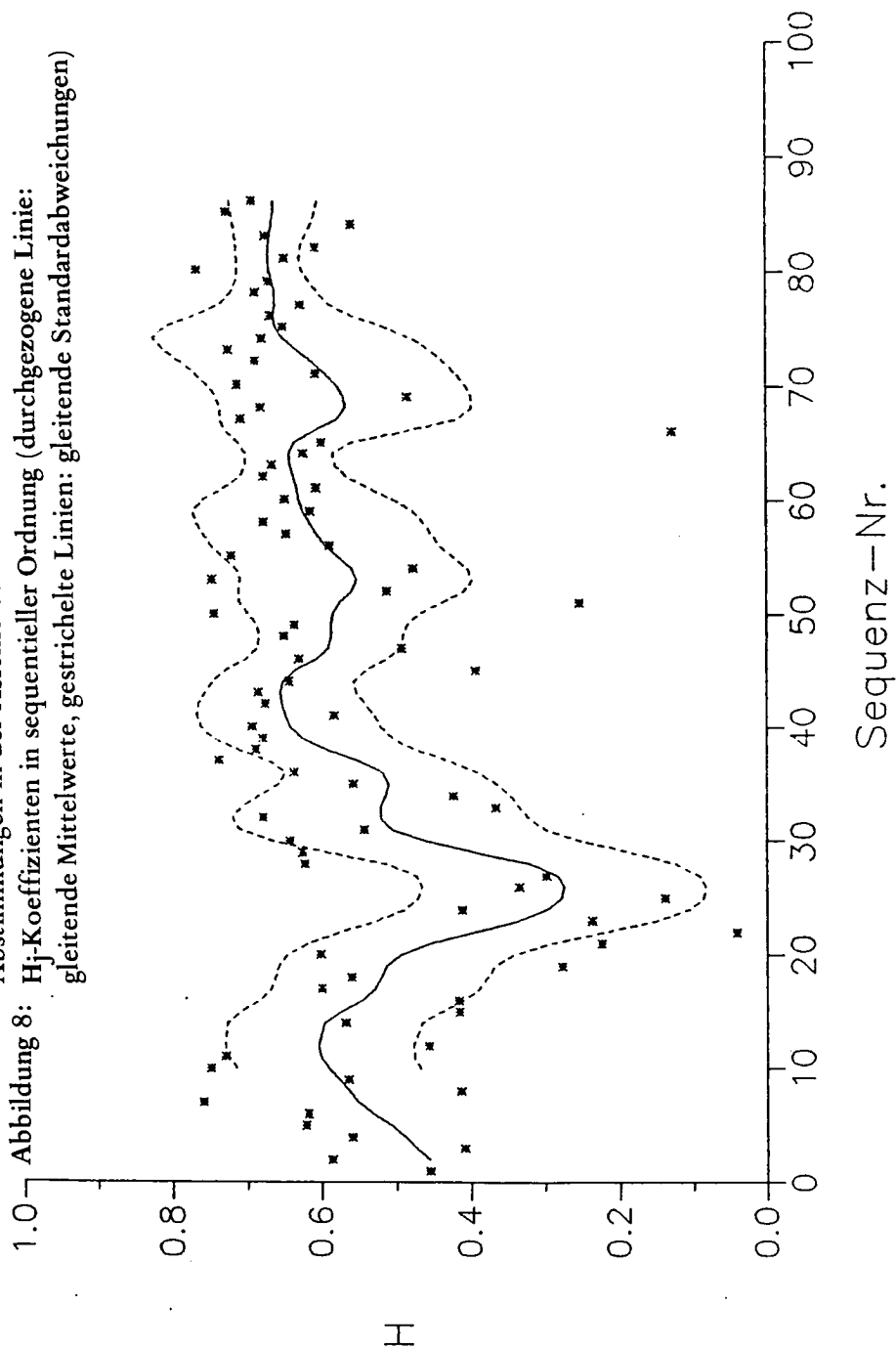


Abbildung 8: H_j-Koeffizienten in sequentieller Ordnung (durchgezogene Linie: gleitende Mittelwerte, gestrichelte Linien: gleitende Standardabweichungen)



hier als ein Lern- und Institutionalisierungsprozeß, als das ungeplante Ergebnis des parlamentarischen Entscheidungshandelns der einzelnen Abgeordneten konzipiert. Tatsächlich widersprach die Formierung kohärenter und relativ disziplinierter Abstimmungsblöcke dem Prinzip der ungebundenen Stimmabgabe, das für die liberale Lehre von der freien Repräsentation konstitutiv ist.³⁹ Andererseits gehen wir davon aus, daß die parlamentarischen Konfliktstrukturen nicht einfach als ungestörte Abbilder gesellschaftlicher Spannungslinien aufgefaßt werden dürfen. Die verfassungsgebenden Versammlungen wurden vor der Entstehung organisierter Massenparteien und der Formierung sozioökonomischer Interessengruppen auf gesamtstaatlicher Ebene gewählt. Tatsächlich hatte der Prozeß der Partei- und Verbandsbildung einen seiner Ausgangspunkte in den verfassungsgebenden Versammlungen selbst. Wenn es also zutrifft, daß die Entstehung der parlamentarischen Konfliktgruppen das Ergebnis von autokatalytischen Lern- und Institutionalisierungsprozessen war, die sich in den Parlamenten vollzogen, müßte sich dies in unseren Daten an der Variation der sequentiell geordneten H_i -Koeffizienten nachweisen lassen. Sie wurden als Homogenitätsmaße eingeführt, in denen die Übereinstimmung des Antwortmustervektors einer einzelnen Vorlage mit einem generellen Abstimmungsmuster zum Ausdruck kommt. Im Grenzfall einer eindimensional-homogenen Konfliktstruktur müßten alle Abstimmungen in einem Parlament eine Skala bilden. Strukturbildung als ein Prozeß, der von einer Situation relativer Strukturlosigkeit ausgeht und zu einem stabilen Konfliktmuster führt, müßte

- (a) in einem Anstieg der H_i -Koeffizienten bis zu einem Sättigungswert nahe 1 und nachfolgender Stabilität der Homogenitätskoeffizienten und/oder
- (b) in einer abnehmenden mittleren Varianz der Homogenitätskoeffizienten bei fortschreitender Sequenzposition resultieren.

Die Abbildungen 7 und 8 zeigen die Positionen der H_i -Koeffizienten als Sterne im Koordinatensystem. Die durchgezogenen Linien sind die Verläufe der exponentiell gewichteten Gleitmittelwerte der H_i -Koeffizienten (geglättet nach dem Modell 4235H).⁴⁰ Die gestrichelten Linien verbinden die Punkte

H_i gegl. \pm SD (10) gegl.

wobei: SD (10) Standardabweichungen der jeweils letzten 10 Punkte sind (gleitende Standardabweichung, Glättung auch hier nach 4235H). Als chronologische Referenz wählen wir die äquidistant geordneten Sequenzpositionen

³⁹ Eberhard Schütt-Wetschky, Grundtypen parlamentarischer Demokratie. Klassisch-altliberaler Typ und Gruppentyp (Freiburg und München 1984).

⁴⁰ Paul Velleman und D. Hoaglin, ABCs of EDA, Duxbury Press, 1981. John W. Tukey, Exploratory Data Analysis, Addison-Wesley, 1977.

nen der Abstimmungen und nicht die realen Zeitabstände, da wir - in Übereinstimmung mit Erkenntnissen amerikanischer Untersuchungen - davon ausgehen, daß Abgeordnete ihr Abstimmungsverhalten an ihren vorausgegangenen Voten, ihrer "voting history" orientieren, weitgehend unabhängig von der Dauer der Zeitintervalle zwischen den Abstimmungen.⁴¹

Die tatsächlichen Eigenschaften der Wertereihen weichen erheblich von den erwarteten Verläufen ab. Im Fall der deutschen Abgeordneten sind die gleitenden Mittelwerte der Homogenitätskoeffizienten annähernd stationär, in Frankreich beobachten wir dagegen zunächst den erwarteten Anstieg, der jedoch etwa mit der zwanzigsten Abstimmung abbricht und sich nach einem tiefen Einbruch erst ab etwa der vierzigsten Abstimmung - bei fortdauernden Schwankungen - im Wertebereich zwischen $H = 0.60$ und $H = 0.80$ einpendelt. Diese Unterschiede zeigen an, daß allgemeine Lerntheorien und der Verweis auf die Eigendynamik von Institutionalisierungsprozessen den tatsächlichen Verlauf der Strukturbildung in beiden Versammlungen nicht hinreichend aufzuklären vermögen. Sie unterschätzen die Bedeutung von Veränderungen des spezifischen Problemhaushalts in beiden Versammlungen im Prozeß der Verfassungsgebung, von Interventionen exogener Entwicklungen und Ereignisse, schließlich den Einfluß der Organisationsform parlamentarischer Konfliktgruppen.

So konnte in der Frankfurter Nationalversammlung bereits Anfang Juni 1848, also rund 14 Tage nach ihrem Zusammentritt von "fertigen, konstituierten und organisierten Parteien" die Rede sein, wobei "Partei" in diesem zeitgenössischen Zitat für den heutigen Begriff der Fraktion steht.⁴² Anfang Oktober 1848 (in der Abbildung 9 etwa Sequenzposition 13) war der Prozeß ihrer programmatischen Abgrenzung und organisatorischen Formierung abgeschlossen. In dieser Phase verteilten sich die Abgeordneten auf bis zu 9 Fraktionen, die ihr Verhalten in Abstimmungen jedoch eng, zum Teil durch fraktionsübergreifende "Kommissionen", koordinierten.⁴³ Ca. 20% der Abgeordneten waren zu diesem Zeitpunkt fraktionslos.⁴⁴ Die Mehrheitsposition hatten die Fraktionen des "Rechten Centrums", dem sich zumeist die Abgeordneten der äußeren Rechten anschlossen. Im Ergebnis bewirkte der Prozeß der Fraktions- und Koalitionsbildung in der Frankfurter Nationalversammlung die Institutionalisierung des rechts-links-Gegensatzes als dominante Konfliktfront. Das rechts-links-Schema bildete das "kognitive Konzept", mit

⁴¹ Aaron Wildavsky, *The Politics of the Budgetary Process* (Boston 1964).

⁴² Karl Biedermann, *Erinnerungen an die Paulskirche* (Leipzig 1849), S. 11.

⁴³ Georg Beseler, *Erlebtes und Erstrebtes 1809- 1859* (Berlin 1884), S. 76.

⁴⁴ Best, *Bildung und Besitz*, S. 326.

dem sich problemlos Abstimmungsvorlagen und Abgeordnete bewerten ließen. Auch die fraktionslosen Mitglieder der Nationalversammlung orientierten ihr Verhalten an diesem Raster und stimmten mehr oder weniger konsistent mit einem der Abstimmungsblöcke.

In unseren Daten kommt der hier skizzierte Prozeß der Institutionalisierung einer Konfliktstruktur in der abnehmenden Streuung der H_1 -Koeffizienten während der Konstituierungsphase der Nationalversammlung zum Ausdruck. Ab September 1848 ordneten sie sich in einem schmalen Wertebereich. Doch hatte diese Konsolidierung keinen Bestand, denn zunehmend geriet die Auseinandersetzung um rechte und linke Positionen, die sich vereinfachend als eine Kontroverse um die soziale Reichweite politischer und ökonomischer Partizipationsrechte deuten läßt, unter den Einfluß einer zweiten Konfliktfront: den großdeutsch-kleindeutschen Gegensatz, in dem es um die territoriale Reichweite einer deutschen Nationalstaatsbildung, die Frage nach der Einbeziehung des österreichischen Kaiserstaats ging.⁴⁵ Abstimmungen, die diesem Konfliktfeld zuzuordnen sind, bilden eine zweite Mokken-Skala, der bei einem Grenzwert von $H > 0.50$ 33 Abstimmungen zugeordnet werden. 28 dieser Abstimmungen sind auch Bestandteil der Hauptskala.

Hier wird deutlich, daß ein Item mehreren eindimensionalen Skalen zugeordnet werden kann. Dementsprechend wird Mokkens Verfahren auch als multiple eindimensionale Skalierung bezeichnet.⁴⁶ Obwohl es demnach zulässig ist, daß sich Skalen, die unterschiedliche Bewertungsdimensionen erfassen, weit überlappen, manifestiert sich in unseren Daten die Pluralisierung der Konfliktstruktur in einer zunehmenden Inhomogenität des Abstimmungsverhaltens: Die steigende Varianz der H_1 -Koeffizienten zwischen der 40. und 55., dann wieder nach der 70. Abstimmung ist dem Umstand zuzuschreiben, daß in diesen Phasen der großdeutsch-kleindeutsche Konflikt auf die Agenda der Nationalversammlung gesetzt wurde. Die Entscheidungen über das Reichsoberhaupt und die Ausdehnung des Reichsgebietes ließen die Frage des Verhältnisses zum österreichischen Kaiserstaat mit seinen umfangreichen nichtdeutschen Gebieten unabweisbar werden. In den Abstimmungen über dieses Thema durchmischten sich dynastische und Staatsloyalitäten mit verfassungspolitischen Grundsatzabwägungen, was erklärt, daß viele der betreffenden Abstimmungen sowohl der links-rechts wie der großdeutsch-klein-

⁴⁵ Ebd., S. 344-349.

⁴⁶ Kees Niemöller u. Wijbrandt van Schuur, "Stochastic Models for Unidimensional Scaling: Mokken and Rasch", in: David McKay et al., *Data Analysis and the Social Sciences* (London 1983), S. 145.

deutschen Skala zugewiesen werden. Erst in den Schlußabstimmungen über die Reichsverfassung gab es dann eine Serie von Voten, die exklusiv der groß-deutsch-kleindeutschen Dimension zugeordnet werden, in denen diese zweite Konfliktfront also "rein" zum Ausdruck kommt. Auch in dieser Phase gab es jedoch Abstimmungen, in denen weiter nach dem Bewertungsschema des links-rechts-Gegensatzes entschieden wurde. Wir beobachten also die komplizierte Situation einer Konfliktstruktur, die kurzfristig zwischen zwei Zuständen wechselte, auf die gleichzeitig zwei unterschiedliche Mechanismen der Strukturbildung einwirkten. Die Kreuzung politischer Konfliktlinien überlastete die institutionellen Strukturen und informellen Prozeduren der Entscheidungsbildung, die sich bis Sommer und Herbst 1848 herausgebildet hatten. In diesem Spannungsfeld zermürbten und zerbrachen schließlich die Gesinnungs- und Handlungsgemeinschaften, die die Grundlage der Fraktionsbildung und der Bindung der Abgeordneten an ihre Mandate gewesen war. Die Spaltung der Nationalversammlung in sezessionierende Teilversammlungen und der vorzeitige Auszug vor allem österreichischer Abgeordneter aus dem Parlament sind Symptome dieses Zerfalls.⁴⁷

Während in Deutschland Fluktuationen der Homogenität des Abstimmungsverhaltens in den Streuungsmaßen zum Ausdruck kommen, wobei die gemittelten H-Koeffizienten annähernd stationär sind, beobachten wir in der Pariser Konstituante ausgeprägte Schwankungen dieser Wertereihe. Der Befund läßt sich plausibel mit der in Frankreich ephemeren Struktur politischer Konfliktgruppen erklären. Zwar waren auch die "Reunionen" in der Pariser Konstituante nach einem rechts-links-Schema differenziert, doch entwickelten sie im gesamten Untersuchungszeitraum keine stabilen Organisationsstrukturen, fluktuierte ihr Anhang, vermochten sie vor allem nicht das Prinzip der gebundenen Stimmabgabe bei ihren Teilnehmern durchzusetzen.⁴⁸ Der links-rechts Gegensatz war hier also nicht in Koalitionen und Kollisionen formal organisierter Konfliktgruppen institutionalisiert, sondern ein eher unscharfes Ordnungsschema, das vor allem durch die Orientierung auf die wechselnden Regierungen der zweiten Republik Kontur erhielt. So fällt der erste "Zyklus" in der Wertereihe der Homogenitätskoeffizienten mit der Etablierung und dem Niedergang der Regierung des Generals Cavaignac zusammen. Während der Phase des Rückgangs der Abstimmungshomogenität formierte sich in den Schlußabstimmungen über die Verfassung zudem noch einmal ein breiter "re-

⁴⁷ Best, *Bildung und Besitz*, S. 264.

⁴⁸ Paul Bastid, *Doctrines et institutions politiques de la Séconde République*, Bd. 1 (Paris 1945), S. 215 ff.

publikanischer" Konsens, der die im Gefolge des Juniaufstandes aufgebrochene Polarisierung zwischen einem linken und einem rechten Lager überbrückte. In der Abstimmung über die Verfassung in ihrer Gesamtheit, die hier mit der Sequenzposition 21 den Tiefpunkt der Wertereihe der Homogenitätskoeffizienten markiert, gab es nur 38 Gegenstimmen von insgesamt 776 Abstimmenden.⁴⁹ Es folgte die Wahl Louis-Napoléon Bonapartes zum Präsidenten und die Installierung des rechtsgerichteten Ministeriums Barrot-Falloux, das durch den "parti de l'ordre", ein aus den Anhängern der verschiedenen Zweige des bourbonischen Königshauses, Bonapartisten und rechten Republikanern zusammengesetztes Bündnis, unterstützt wurde. Während es General Cavaignac nicht gelungen war, einen stabilen parlamentarischen Rückhalt zu gewinnen, formierte sich nun ein gouvernementales Lager, dem eine linke Opposition gegenüberstand. Damit gewann der links-rechts-Gegensatz als handlungssteuernde Konfliktfront an Prägnanz und Stabilität. Der Fraktionszwang, der in der Nationalversammlung Konformität des Abstimmungsverhaltens bewirkte, hatte in Frankreich in den Interventionen der Regierung, vor allem durch Patronage und Klientelismus, ein Äquivalent.⁵⁰ Diese Beobachtung sollte jedoch nicht zu der Schlußfolgerung verleiten, es habe zugleich auch eine Stabilisierung der Mehrheiten gegeben: Das gouvernementale Lager und die linke Opposition blieben an ihren Rändern ephemere und inhomogene. Die immer noch beachtlichen Fluktuationen der Homogenitätskoeffizienten und ihr erheblicher Streubereich auch nach der Inauguration des Präsidenten Bonaparte am 20. Dezember 1848 (zwischen Sequenzposition 29 und 30) zeigen dies an.

Festzuhalten bleibt, daß der zeitliche Ablauf der Strukturbildung in beiden Parlamenten in der Tat als ein Sozialisations- und Institutionalisierungsprozeß beschrieben werden kann, in dessen Verlauf sich diffuse und inhomogene Orientierungen verstetigen und institutionell verfestigen. Offensichtlich begünstigt die Logik der Entscheidungsbildung in Parlamenten die Formierung "einfacher" Strukturen, in denen sich Issues in niederdimensionalen Konfliktträumen anordnen und sich Abgeordnete in polarisierten Konfliktgruppen gegenüber treten. Diese Deutung wird auch durch den Umstand gestützt, daß das in beiden Konstituanten beobachtete Konfliktmuster in sehr ähnlicher Weise in anderen "Parlamenten der ersten Stunde" auftrat: Im Kongress der Vereinigten Staaten, dem französischen Konvent von 1792 oder der mexikani-

⁴⁹ Abstimmungsergebnis in Heinrich Best, Die Abgeordneten der Assemblée Nationale Constituante. Sozialprofil und legislatives Verhalten (Köln 1983), S. 97.

⁵⁰ Roger Price, The French Second Republic. A Social History (London 1972), S. 225-245.

schen Konstituante von 1917.⁵¹ Andererseits hat aber die Rekonstruktion der Dynamik des Prozesses der Strukturbildung gezeigt, daß es eine unzulässige Verkürzung wäre, Parlamente ausschließlich als Interaktionssysteme zu modellieren, deren Strukturbildung "selbstreferentiell" dem Erfordernis folgt, die Komplexität von Problemlagen zu verringern und Konfliktgruppen zu stabilisieren. Gerade der Fall Deutschland liefert mit dem großdeutsch-kleindeutschen Konflikt das Beispiel dafür, daß auch äußere Problemlagen intervenieren und einen Grad an "Sperrigkeit" erreichen können, der sich jedem Versuch zur Realitätsvereinfachung widersetzt. Die Formierung einer zweiten dominanten Konfliktfront um die territoriale Reichweite der Nationalstaatsbildung, die man unter normativen Gesichtspunkten vielleicht als erwünschte Antwortbereitschaft eines Parlaments auf die Pluralität der vielfach parzellierten politischen Landschaft Deutschlands bewerten könnte, erwies sich tatsächlich als ein Desaster für den Parlamentarisierungsprozeß: Im Spannungsfeld sich kreuzender Konfliktlinien zermürbten und zerbrachen schließlich jene parlamentarischen Gesinnungs- und Handlungsgemeinschaften, die vielleicht in der Lage gewesen wären, das politische Programm der Reichsverfassung gegenüber den traditionellen Machttägern zu verwirklichen.

Die Möglichkeit, von den Schwankungen und der Streuung der Koeffizientenreihen auf bestimmte historische Ereignisse und institutionelle Entwicklungen rückzuschließen, bestätigt noch einmal die Angemessenheit der Mokken-Skalierung für die serielle Analyse namentlicher Abstimmungen: Mit ihr lassen sich nicht nur die Ergebnisse, sondern auch die Dynamiken von Strukturbildungsprozessen in Parlamenten und anderen Entscheidungskörperschaften abbilden. Für einen in der historischen Forschung weit verbreiteten Datentyp steht nunmehr ein mächtiges Verfahren bereit, das frei ist von den Nachteilen der sonst in diesem Forschungskontext eingesetzten Methoden.

⁵¹ Peter H. Smith, "The making of the Mexican Constitution", in: William O. Aydelotte (ed.), *The History of Parliamentary Behavior* (Princeton/N.J. 1977), S.186-224; Alison Patrick, *The Men of the First French Republic: Political Alignments in the National Convention of 1792* (Baltimore 1987); John Hoadley, "The Emergence of Political Parties in Congress, 1789-1803", in: *American Political Science Review*, 74 (1980), S.757-779.